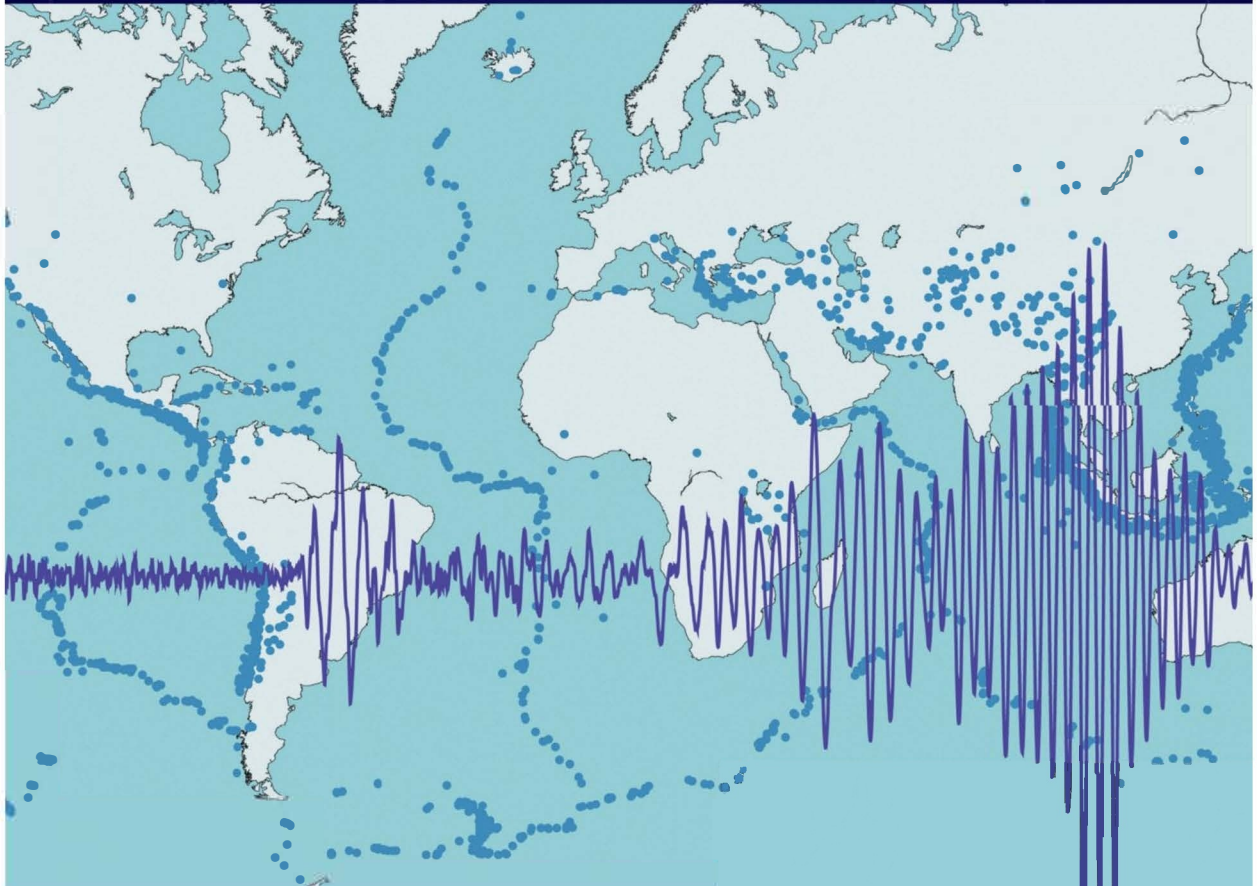


PRINCIPLES OF SEISMOLOGY

SECOND EDITION

Agustín Udías and Elisa Bufo



Principles of Seismology

Second Edition

The second edition of *Principles of Seismology* has been extensively revised and updated to present a modern approach to observation seismology and the theory behind digital seismograms. It includes: a new chapter on earthquakes, Earth's structure and dynamics; a considerably revised chapter on instrumentation, with new material on processing of modern digital seismograms and a list of website hosting data and seismological software; and 100 end-of-chapter problems. The fundamental physical concepts on which seismic theory is based are explained in full detail with step-by-step development of the mathematical derivations, demonstrating the relationship between motions recorded in digital seismograms and the mechanics of deformable bodies. With chapter introductions and summaries, numerous examples, newly drafted illustrations and new color figures, and an updated bibliography and reference list, this intermediate-level textbook is designed to help students develop the skills to tackle real research problems.

Agustín Udías is Emeritus Professor at the Universidad Complutense, Madrid, Spain. He is the author of many papers about seismicity, seismotectonics of the Azores-Gibraltar, and the physics of seismic sources, and has also written several textbooks. Professor Udías has served as Editor-in-Chief of *Física de la Tierra* and the *Journal of Seismology*, and as the Vice-President of the European Seismological Commission. He is a member of the *Accademia Europeae*, the Seismological Society of America, the American Geophysical Union, amongst other societies, and is a Fellow of the Royal Astronomical Society.

Elisa Buforn is Professor of Geophysics at the Universidad Complutense, Madrid, Spain, where she teaches courses on geophysics, seismology, physics, and numerical methods. Professor Buforn's research focuses on source fracture processes, seismicity, and seismotectonics and she is Editor-in-Chief of *Física de la Tierra* and on the Editorial Board of the *Journal of Seismology*. She is a member of the Seismological Society of America, American Geophysical Union, Royal Astronomical Society, and Real Sociedad Española de Física, amongst other societies.

Principles of Seismology

Second Edition

AGUSTÍN UDÍAS

Universidad Complutense, Madrid

ELISA BUFORN

Universidad Complutense, Madrid



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107138698

DOI: [10.1017/9781316481615](https://doi.org/10.1017/9781316481615)

© Agustín Udías and Elisa Buform 2018

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2018

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Udías Vallina, Agustín, author. | Buform, Elisa, 1954– author.

Title: Principles of seismology / Agustín Udías, Universidad Complutense, Madrid, Elisa Buform, Universidad Complutense, Madrid.

Description: Second edition. | Cambridge : Cambridge University Press, 2017. | Includes bibliographical references and index.

Identifiers: LCCN 2017024371 | ISBN 9781107138698

Subjects: LCSH: Seismology. | Seismology – Mathematics. | Wave mechanics – Mathematics.

Classification: LCC QE534.2 .U35 2017 | DDC 551.22–dc23

LC record available at <https://lcn.loc.gov/2017024371>

ISBN 978-1-107-13869-8 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

<i>Preface and Acknowledgments</i>	<i>page xv</i>
1 Seismology, the science of earthquakes	1
1.1 The historical development	1
1.2 Seismology a multidisciplinary science	4
1.3 Divisions of seismology	5
1.4 Theory and observations	6
1.5 International cooperation	8
1.6 Books, journals, and websites	9
1.7 Summary	12
2 Earthquakes, Earth structure, and dynamics	14
2.1 Earthquakes: natural disasters	14
2.2 Size of earthquakes	17
2.3 Earthquakes and faults	18
2.4 Spatial distribution of earthquakes	20
2.5 Temporal distribution of earthquakes	22
2.6 Earth's structure: crust, mantle, and core	23
2.7 Plate tectonics	26
2.8 Earthquake risk, prediction, and prevention	29
2.9 Summary	30
3 Instrumentation and digital data processing	32
3.1 The historical evolution of seismographs	32
3.2 The theory of the seismometer	35
3.3 Recording systems, magnification, and dynamic range	39
3.4 Electromagnetic seismographs	40
3.5 Digital seismographs	45
3.6 Processing digital seismograms	47
3.6.1 Data acquisition	47
3.6.2 Radial and transverse components	48
3.6.3 Removing the instrumental response	49
3.6.4 Spectral analysis	50
3.6.5 Filtering seismograms	51

3.7 Accelerographs	54
3.8 Other types of seismologic instruments	55
3.9 Summary	58
3.10 Problems	58
4 Basic concepts and equations of an elastic medium	60
4.1 Displacement, strain, and stress	60
4.1.1 Eigenvalues and eigenvectors	64
4.2 Elasticity coefficients	65
4.3 The influence of temperature	68
4.4 Work, energy, and heat considerations	70
4.5 Equations of continuity and motion	72
4.5.1 The equations of continuity of mass and energy	72
4.5.2 The equation of motion or momentum	73
4.6 The Lagrangian formulation	77
4.7 Potential functions of displacements and forces	78
4.8 The Green and Somigliana functions of elastodynamics	79
4.9 Theorems of reciprocity and representation	81
4.10 Summary	84
4.11 Problems	85
5 Waves in an infinite elastic medium	87
5.1 Wave equations for an elastic medium	87
5.2 Solutions of the wave equation	89
5.2.1 Wave fronts and rays	91
5.2.2 Waves of several frequencies	93
5.3 Displacement, velocity, and acceleration	93
5.4 The propagation of energy	95
5.4.1 Phase and group velocities	96
5.5 The effect of gravity on wave propagation	97
5.6 Plane waves	98
5.7 The geometry of P and S wave displacements	101
5.8 Particular forms of the potentials	103
5.9 Spherical waves	105
5.10 Cylindrical waves	109
5.11 Summary	112
5.12 Problems	112
6 Reflection and refraction	115
6.1 Snell's law	115
6.2 Reflection and refraction in two liquid media	116
6.2.1 Normal incidence	120
6.2.2 Critical incidence	120
6.2.3 Inhomogeneous waves	121

6.2.4 Reflected and transmitted energy	122
6.2.5 Reflection on a free surface	126
6.3 Reflection and refraction in elastic media	127
6.3.1 Incident SH waves	127
6.3.2 Critical incidence and inhomogeneous waves	129
6.3.3 Incident P and SV waves	130
6.4 Reflection on a free surface	132
6.4.1 Incident SH waves	133
6.4.2 Incident P waves	133
6.4.3 Incident SV waves	135
6.4.4 Critical reflection of SV waves	136
6.4.5 The partition of energy	136
6.5 Motion at the free surface	139
6.5.1 Incident P waves	139
6.5.2 Incident S waves	140
6.5.3 Apparent angles of incidence and polarization	141
6.6 Summary	143
6.7 Problems	143
7 Body wave propagation in layered media	146
7.1 Wave propagation in the (x, z) plane	146
7.2 The equation for the displacement–stress vector	148
7.3 The propagator matrix	150
7.4 A layered medium with constant parameters	152
7.4.1 Eigenvalues and eigenvectors	153
7.4.2 The propagator matrix for SH motion	154
7.5 SH motion in an elastic layer over a half-space	155
7.6 The general problem	157
7.6.1 SH motion	158
7.6.2 P–SV motion	158
7.7 Summary	159
7.8 Problems	159
8 Ray theory. Media of constant velocity	161
8.1 The eikonal equation	161
8.1.1 The condition of validity	164
8.2 Ray trajectories	165
8.3 Ray trajectories and travel times. A homogeneous half-space	167
8.4 A layer over a half-space with constant velocities	171
8.5 The dipping layer	176
8.6 A plane-layered medium	178
8.7 Summary	181
8.8 Problems	181

9 Ray theory. Media of variable velocity	183
9.1 A variable velocity with depth	183
9.2 The Lagrangian formulation	187
9.3 The change of distance with the ray parameter	188
9.4 The velocity distribution with ζ constant	190
9.5 A linear increase of velocity with depth	191
9.6 Distributions of velocity with depth	194
9.6.1 A gradual increase of velocity	194
9.6.2 A rapid increase in velocity	195
9.6.3 A decrease of velocity. A low-velocity layer	198
9.7 Travel times for deep foci	200
9.8 Reflected rays	200
9.9 Determination of the velocity distribution	202
9.10 The energy propagated by ray beams. Geometrical spreading	204
9.11 Summary	206
9.12 Problems	207
10 Ray theory. Spherical media	209
10.1 The geometry of ray trajectories and displacements	209
10.2 A sphere of constant velocity	212
10.3 A sphere with a velocity that is variable with the radius	213
10.3.1 The change of distance with the ray parameter	215
10.4 A velocity distribution with ζ constant	217
10.5 Rays with circular trajectory	217
10.6 Distribution of the velocity with the radius	218
10.6.1 A normal distribution	219
10.6.2 A rapid increase in velocity	220
10.6.3 A decrease in velocity. A low-velocity layer	221
10.7 Determination of velocity distribution	222
10.8 Energy propagation by ray beams. Geometric spreading	224
10.9 Summary	226
10.10 Problems	226
11 Travel times and the structure of the Earth	228
11.1 Observations and methods	228
11.1.1 Refraction and wide-angle reflection	230
11.1.2 Vertical reflection	231
11.1.3 Seismic tomography	233
11.1.4 Receiver functions	235
11.2 Distribution of velocity, elasticity coefficients, and density	236
11.3 The crust	238
11.3.1 The mineralogical composition of the crust	243
11.4 The upper mantle	244

11.4.1 The lithosphere and the asthenosphere	245
11.5 The lower mantle	247
11.6 The core	251
11.7 Summary	257
11.8 Problems	258
12 Surface waves	259
12.1 Rayleigh waves in a half-space	259
12.1.1 Displacements of Rayleigh waves	263
12.2 A liquid layer over a rigid half-space. Guided waves	265
12.2.1 Constructive interference	266
12.2.2 The dispersion equation and curves	267
12.2.3 Displacements	268
12.3 An elastic layer over a half-space. Love waves	269
12.3.1 Constructive interference	271
12.3.2 Dispersion curves	272
12.3.3 Displacements	274
12.4 An elastic layer over a half-space. Rayleigh waves	276
12.5 Surface waves in layered media	279
12.5.1 Love waves in a layer over a half-space	279
12.6 Surface waves in a spherical medium	280
12.7 Stoneley waves	281
12.8 Summary	283
12.9 Problems	284
13 Wave dispersion. Phase and group velocities	285
13.1 Phase and group velocities	285
13.2 Groups of waves	287
13.3 The principle of a stationary phase	289
13.4 Characteristics of dispersed waves	292
13.5 Determination of group and phase velocities. Instantaneous frequencies	294
13.5.1 The group velocity	294
13.5.2 The phase velocity	295
13.6 Determination of phase and group velocities. Fourier analysis	297
13.6.1 Fourier analysis of seismograms	297
13.6.2 The phase velocity	298
13.6.3 The group velocity	298
13.7 Dispersion curves and the Earth's structure	300
13.7.1 Observations	300
13.7.2 Interpretation	302
13.8 Summary	306
13.9 Problems	306

14 Free oscillations of the Earth. Theory and observations	308
14.1 Standing waves and modes of vibration	308
14.2 Vibrations of an elastic string of finite length	310
14.3 Vibrations of an elastic rod	313
14.3.1 Longitudinal vibrations	313
14.3.2 Torsional vibrations	315
14.4 The general problem. The Sturm–Liouville equation	317
14.5 Free oscillations of a homogeneous liquid sphere	318
14.6 Free oscillations of an elastic sphere	322
14.7 Toroidal modes	324
14.8 Spheroidal modes	327
14.9 Effects on free oscillations	329
14.10 Observations	331
14.11 Summary	333
14.12 Problems	334
15 Anelasticity and anisotropy	335
15.1 Anelasticity and damping	335
15.1.1 Anelasticity	336
15.1.2 Harmonic excitation of a Maxwellian body	337
15.1.3 Damped harmonic motion. The Q coefficient	338
15.2 Wave attenuation	341
15.3 The attenuation of body and surface waves	343
15.3.1 Body waves	343
15.3.2 Surface waves	345
15.4 The attenuation of free oscillations	346
15.5 The attenuation of coda waves	348
15.6 Attenuation in the Earth	349
15.7 Anisotropy	351
15.8 Wave propagation in anisotropic media	352
15.8.1 Body waves	352
15.8.2 Surface waves	354
15.9 Anisotropy in the Earth	355
15.10 Summary	356
15.11 Problems	357
16 Focal parameters of earthquakes	358
16.1 Location of an earthquake focus	358
16.1.1 Macroseismic determination of epicenter locations	359
16.1.2 Instrumental determination. Graphical methods	360
16.1.3 Numerical methods	362
16.2 Joint hypocenter determination	364

16.3	Seismic intensity	365
16.3.1	Isoseismal or intensity maps	367
16.4	Magnitude	368
16.4.1	Scales of magnitude	368
16.4.2	The saturation of magnitude scales	372
16.5	Seismic energy	373
16.6	The seismic moment, stress drop, and average stress	374
16.7	Summary	378
16.8	Problems	379
17	Basic theory of earthquake mechanism	380
17.1	Earthquakes and faults	380
17.2	Equivalent forces. Point sources	382
17.2.1	Formulation using Green's function	384
17.2.2	Single and double couples	385
17.3	Fractures and dislocations	388
17.4	The Green function for an infinite medium	390
17.4.1	The radial force	390
17.4.2	An impulsive force in an arbitrary direction	391
17.5	The separation of near and far fields	396
17.5.1	The near field	397
17.5.2	The far field	397
17.6	A shear dislocation or fracture. The point source	400
17.6.1	The radiation pattern	401
17.6.2	The geometry of a shear fracture	404
17.7	The source time function	406
17.8	The equivalence between forces and dislocations	408
17.9	Summary	411
17.10	Problems	412
18	The seismic moment tensor	414
18.1	Definition of the moment tensor	414
18.2	The moment tensor and elastic dislocations	418
18.2.1	An explosive source	418
18.2.2	Shear fracture	419
18.3	Eigenvalues and eigenvectors	420
18.4	Types of sources and separation of the moment tensor	422
18.5	Displacements due to a point source	424
18.6	The temporal dependence	425
18.7	The centroid moment tensor	427
18.8	Inversion of the moment tensor	427
18.9	Summary	429
18.10	Problems	429

19 Simple models of fracture	431
19.1 Source dimensions. Kinematic models	432
19.2 Rectangular faults. Haskell's model	434
19.3 Circular faults. Brune's model	438
19.4 Nucleation, propagation, and arrest of a rupture	440
19.5 Dynamic models of fracture	444
19.5.1 The static problem	444
19.5.2 The dynamic problem	445
19.6 Friction models of fracture	448
19.7 The complexity of a fracture	449
19.7.1 The cohesive zone	449
19.7.2 Barriers and asperities	451
19.7.3 Acceleration spectra	453
19.8 Summary	454
19.9 Problems	455
20 Methods of determination of source mechanisms	456
20.1 Parameters and observations	456
20.2 P wave first motion polarities. Fault plane solutions	457
20.2.1 Graphical methods	460
20.2.2 Numerical methods	461
20.3 Wave-form modeling	463
20.4 Inversion of the moment tensor	468
20.4.1 The inversion of Rayleigh waves	468
20.5 Amplitude spectra of seismic waves	470
20.6 Determination of the slip distribution over the fault plane	471
20.7 Summary	473
20.8 Problems	475
21 Seismicity, seismotectonics, seismic risk, and prediction	477
21.1 The spatial distribution of earthquakes	477
21.2 The temporal distribution of earthquakes	480
21.3 Seismic cycles	481
21.4 The distribution of magnitudes	483
21.5 Models of the occurrence of earthquakes	485
21.6 Seismotectonics	487
21.7 Seismic hazard and risk	490
21.8 The prediction of earthquakes	493
21.9 Summary	496
<i>Appendix 1 Vectors and tensors</i>	497
A1.1 Definitions	497
A1.2 Operations with vectors and tensors	498

A1.3 Vector and tensor calculus	500
<i>Appendix 2 Cylindrical and spherical coordinates</i>	502
A2.1 Cylindrical coordinates	502
A2.2 Spherical coordinates	504
<i>Appendix 3 Bessel and Legendre functions</i>	506
A3.1 Bessel functions	506
A3.2 Spherical Bessel functions	507
A3.3 Legendre functions	508
A3.4 Associate Legendre functions	509
<i>Appendix 4 Fourier and Laplace transforms</i>	511
A4.1 Periodic functions	511
A4.2 Non-periodic functions	512
A4.3 Convolution and correlation	513
A4.4 Sampled functions of finite duration	514
A4.5 Laplace transform	516
<i>Appendix 5 Parameters of the Earth</i>	518
<i>Appendix 6 The interior of the Earth</i>	519
<i>Appendix 7 Important earthquakes</i>	520
<i>Bibliography</i>	522
<i>References</i>	526
<i>Index</i>	540

Color plates are to be found between pp. 296 and 297.

Preface and Acknowledgments

A second edition of this textbook, first published 17 years ago, is a wonderful opportunity to review its contents and improve its pedagogical orientation, in view of the many comments and interactions received, teaching experience, and experimental progress in seismology over the intervening years. Seismology is the science of earthquakes, which are both natural disasters profoundly affecting human lives, and a subject of study through application of the principles of the physical sciences. These two aspects are linked, since an important aim of the study of seismology is to mitigate the terrible effects of earthquakes through a more complete knowledge of their nature. Seismology provides us also with a powerful instrument to study the constitution and dynamics of the Earth. To emphasize these different aspects of seismology, [Chapter 2](#) has been added to present, as an introduction, the complex phenomenon of earthquakes from a narrative point of view. As a physical science, the fundamentals of seismology are based on analysis of the seismic waves produced by earthquakes and registered by seismographs. The importance of this aspect is shown by presenting the analysis of seismographic digital data in [Chapter 3](#), so that it can be used in subsequent chapters. This is a unique feature not present in other texts on seismology. Thus, this approach has been used in the new edition. The text is at an introductory level for students in the last years of the European Licentiate and the first year of Masters programs or American upper-division undergraduate courses and first graduate courses, and at similar levels of study in other countries. As a first book, no previous knowledge of seismology, as such, is assumed of the student. The book's emphasis, as indicated by its title, is on the fundamental physical principles which constitute the basis of the analysis of seismic waves and their basic development. In consequence, a number of topics have been selected. It has been noticed that sometimes even graduate students lack a true grasp of the fundamental physical principles underlying some aspects of seismology. In this book, the most fundamental concepts are, therefore, developed in detail, with their mathematical developments fully worked out. Simple cases, such as one-dimensional problems and those in liquid media, are used as introductory topics. In some instances, more difficult subjects are introduced, although not fully developed. In these cases references to more advanced books and articles are given where they can be found. In each chapter, problems are proposed, some of them are fully solved in the electronic material. As an innovation in this edition, for some of the problems seismogram digital data are used, which are given in the electronic material. Details of websites from where data and programs can be retrieved are provided. The reader can access the electronic material at www.cambridge.org/UdiasBufom, and it is referenced in the text as EM.

The book presupposes a certain level of knowledge of mathematics and physics. Knowledge of mathematics at the level of calculus and ordinary and partial differential equations, as well as a certain facility for vector and tensor analysis, are assumed. Cartesian,

spherical, and cylindrical coordinates, and some functions such as Legendre and Bessel functions are used. Tensor index notation is used preferentially throughout the book. Fundamental ideas about certain mathematical subjects are given briefly in [Appendixes 1 to 4](#). Basic knowledge of the mechanics of a continuous medium and of the theory of elasticity is also presupposed, but the reader is reminded about the basic equations of elasticity in [Chapter 4](#) and for other topics is referred to textbooks on elasticity that are cited in the Bibliography.

Throughout the book there is an emphasis on the fundamental theoretical aspects of seismology and observations are treated briefly. Thus, some readers will miss discussion of recent results; we refer them to the excellent books by Lay and Wallace (1995) and Stein and Wysession (2003). Also, more advanced developments of the theory of wave propagation and generation are not treated; see, for example, Aki and Richards (1980), Ben Menahem and Singh (1981), and Dahlen and Tromp (1998). We hope that our book is a good introduction to these excellent advanced books. It is difficult to decide where to stop with subjects treated in a textbook that is designed as an introduction. We have selected to develop only, but with all mathematical detail, the very basic problems. In this sense, as was mentioned in the preface of the first edition, this book is different from those that already exist. The style and approach are also sometimes different, and reflect those of the authors.

After the introductory two chapters providing a short narrative presentation of the phenomenon of earthquakes, [Chapter 3](#) gives the theory of seismographs and the analysis of seismograms in digital form. In this way digital seismograms can be used in subsequent chapters and problems thereby included. The following chapters are dedicated to the fundamentals of elasticity theory ([Chapter 4](#)), solutions of the wave equation ([Chapter 5](#)), the propagation of body waves ([Chapters 6 and 7](#)), ray theory ([Chapters 8 to 11](#)), and surface waves ([Chapters 12 and 13](#)), normal modes and free oscillations ([Chapter 14](#)), with an introduction to anelasticity and anisotropy ([Chapter 15](#)). Five chapters are devoted to the study of the earthquake source and the focal mechanism ([Chapters 16 to 20](#)). The final one ([Chapter 21](#)) introduces the reader to the problems of seismicity, seismotectonics, and seismic risk. [Appendixes 1 to 4](#) cover some mathematical tools, [Appendixes 5 to 7](#) give some helpful information. The Bibliography includes books on seismology and related topics. Other references cited in the text are given separately. Some books are listed as references, so one must use both lists.

The authors wish to thank in the first place all of our students over many years at the Universidad Complutense in Madrid, to whom we are indebted for their questions and suggestions, which have helped us to write this second edition, and for their patience during our lectures. We must thank also a long list of seismologists, some of them former students, who will be difficult to name without omissions, and we hope, therefore, that they will all feel included in our thanks. We thank IRIS (Incorporated Research Institutions for Seismology) for providing some of the digital seismograms used in examples and problems. Finally, we very much appreciate Cambridge University Press for offering to prepare this new edition.

The term *seismology* is derived from two Greek words, *seismos*, shaking, and *logos*, science or treatise. Earthquakes were called *seismos gês* in Greek, literally, shaking of the Earth; the Latin term is *terrae motus*, and from the equivalents of these two terms come the words used in most occidental languages. Seismology means, then, the science of the shaking of the Earth or the *science of earthquakes*. The term seismology and similar ones in other occidental languages (séismologie, sismología, Seismologie, etc.) began to be used at around the middle of the nineteenth century. In this chapter we present a very short overview of the history of seismology (brief information about pertinent historical developments can also be found in each chapter), a discussion of seismology considered as a multidisciplinary science, its theoretical and observational aspects, international cooperation, and a summary of books, journals, and websites.

1.1 The historical development

In antiquity, the first rational explanations of earthquakes, beyond mythical stories, are from Greek natural philosophers beginning with Thales of Miletos in the sixth century BC. Aristotle (in the fourth century BC) discussed the nature and origin of earthquakes in the second book of his treatise on meteors (*Meteorologica*). The term meteors was used by the ancient Greeks for a variety of phenomena believed to take place somewhere above the Earth's surface and below the orbit of the Moon, such as rain, wind, thunder, lightning, comets, but also earthquakes and volcanic eruptions inside the Earth. The term meteorology derives from this word, but in modern use it refers only to atmospheric phenomena. Aristotle, following other Greek authors, such as Anaxagoras, Empedocles, and Democritus, proposed that the cause of earthquakes consists in the shaking of the Earth due to underground dry heated vapors or winds trapped in its interior and trying to leave toward the exterior. This explanation was part of his general theory for all meteors caused by various types of exhalations of gas or vapor (*anathymiaseis*) that extend from inside the Earth to the Lunar orbit. This theory was spread more widely by the encyclopedic Roman authors Lucius Anneus Seneca and Gaius Plinius (Pliny the Elder). It was commented upon by medieval philosophers such as Albert the Great and Thomas of Aquinas, and, with small changes, was accepted in the West until the seventeenth century. For example, in 1678 Athanasius Kircher, a Jesuit professor at the Roman College, in his book *Mundus Subterraneus*, related earthquakes and volcanoes to a system of fire conduits (*pyrophylacia*) and another of air (*aerophylacia*) inside the Earth. With the birth of modern science in the

seventeenth and eighteenth centuries, Martin Lister and Nicolas Lemery proposed that earthquakes are caused by explosions of flammable material concentrated in some interior regions. This explanation was accepted and propagated by Isaac Newton and Georges-Louis Buffon.

The great Lisbon earthquake of 1 November 1755, which caused widespread destruction in that city and produced a large tsunami, may be considered the starting point of modern seismology. In 1758, Joachim Moreira de Mendonça published *Historia Universal dos Terremotos*, a study on the causes of earthquakes with one of the first global catalogs (Fig. 1.1). In 1760, John Michell was one of the first to relate the shaking due to earthquakes to the propagation of elastic waves inside the Earth. This idea was further developed by, among others, Thomas Young, Robert Mallet, and John Milne. Descriptions of damage due to earthquakes and the compilation of lists of their occurrence can be traced back to very early dates. Sometimes these lists include other natural disasters such as floods, famines, and plagues. The first catalogs of earthquakes published in Europe were at the end of the seventeenth century, by Marcello Bonito and Joannes Zahn (Chapters 2 and 16).

Robert Mallet was aware of starting a new science when in 1848 he wrote: “The present paper constitutes, so far as I am aware, the first attempt to bring the phenomenon of the earthquake within the range of exact science, by reducing to system the enormous mass of disconnected and often discordant and ill observed facts which the multiplied narratives of earthquakes present, and educing from these, by appeal to the established laws of the higher mechanics, a theory of earthquake motion” (Mallet 1848). His study of the Naples earthquake of 1857 constitutes one of the first basic works of modern seismology (Mallet, 1862). Mallet developed the theory of the seismic focus from which elastic waves are propagated in all directions and connected the occurrence of earthquakes with changes in the Earth’s crust that are often attended by dislocations and fractures, abandoning the explosive theory. Geologists such as Charles Lyell and Eduard Suess had related earthquakes to volcanic and tectonic motions, and, at the beginning of the twentieth century, Ferdinand Montessus de Ballore and August Sieberg assigned the cause of earthquakes to orogenic processes and contributed to many aspects of observational seismology (Chapter 16 and 17). Two fundamental steps in the process of quantification and formalization of seismology are the development of seismic instrumentation to record the ground motion produced by earthquakes, opening the possibility of having quantified observations and application of the principles of the theory of continuous mechanics. Thus, seismology ceased to be a purely naturalist science. The first instruments used to observe the shaking of the ground were based on oscillations of a pendulum and started to be used in around 1830. By the end of the century, the first seismographic continuous recordings had been produced. In 1889, Ernst von Rebeur Paschwitz recorded, in Potsdam, an earthquake that took place in Tokyo; this was the first seismogram recorded at a large distance. Among the first names related to the development of seismologic instrumentation are those of John Milne and Fusakishi Omori, with the inclined pendulum, Emil Wiechert with the inverted pendulum, Boris B. Galitzin with the electromagnetic seismograph, and Hugo Benioff with variable magnetic reluctance (Chapter 3). Towards the end of the nineteenth century and beginning of the twentieth, Wiechert, Karl Zöppritz, and Richard D. Oldham, among others, published the first studies of the

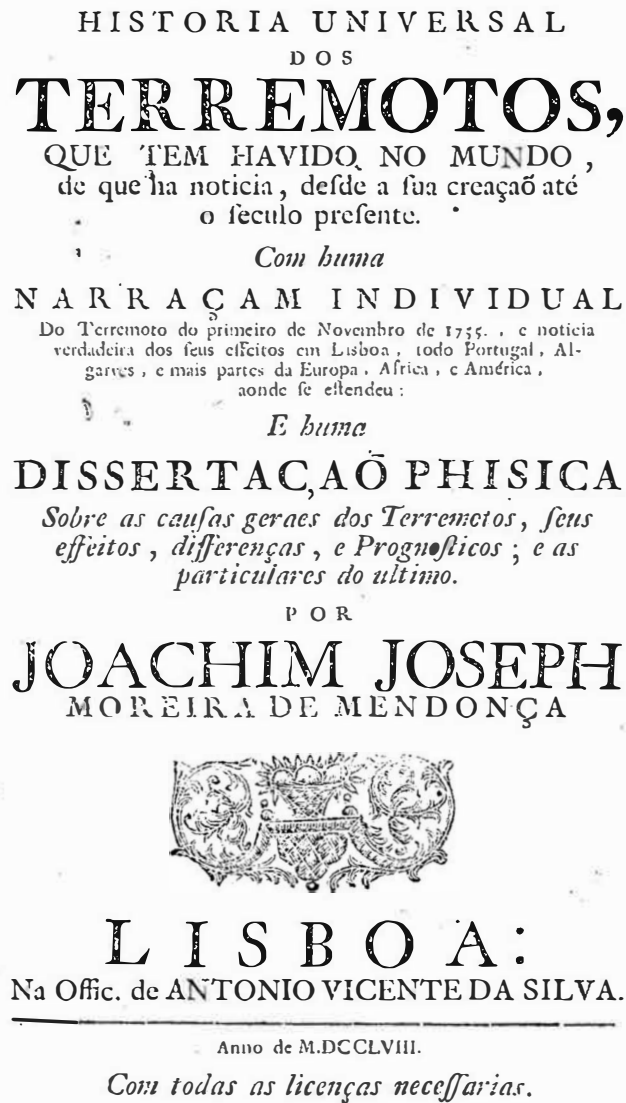


Fig. 1.1

Title page of Joachim Moreira de Mendonça (1758) *Historia Universal dos Terremotos*, written after the 1755 Lisbon earthquake (New York Public Library).

propagation of seismic waves, based on early work on the theory of elasticity (Chapter 4). The first models of the interior of the Earth based on seismologic observations were proposed between 1900 and 1940 by, among others, Oldham, Beno Gutenberg, Harold Jeffreys, Keith E. Bullen, and James B. Macelwane (Chapter 9).

Since 1945, seismology has experienced a very rapid development. Details of this development and names associated with it can be found in the introductions to each chapter and elsewhere in this book. In this rapid evolution, two important subjects are the propagation of elastic waves in the Earth and the mechanism of the generation of earthquakes. Both include theoretical and observational aspects. Observations of seismic waves have improved and multiplied with the development of seismologic instrumentation, from early mechanical seismographs with analog recording, to the present systems with a broadband response, electronic amplification, and digital recording ([Chapter 3](#)). In the study of the propagation of seismic waves, the Earth is approximated by models that have progressed from the early very simple models of homogeneous elastic media or media divided into layers to those with three-dimensional heterogeneity, including anelasticity and anisotropy ([Chapters 4 to 15](#)). Models of the source of earthquakes have developed from simple models of point foci to those including the complex process of fracture and friction of crustal material along faults ([Chapters 16 to 20](#)). These developments have contributed to an increase in knowledge about the complex processes that cause earthquakes and the properties and composition of the materials of the Earth's interior. Other aspects of seismology concern the occurrence of earthquakes in time and space (seismicity), its relation to tectonic processes (seismotectonics), the evaluation of seismic hazard and risk, and the problem of earthquake prediction and earthquake early-warning systems. These aspects of seismology have also significantly expanded in the latest years ([Chapter 21](#)).

1.2 Seismology, a multidisciplinary science

Recent trends in seismology tend to overemphasize those aspects related to the generation and propagation of seismic waves. With this emphasis, Keiti Aki and Paul G. Richards (1980) define seismology as a science based on data called seismograms, which are the records of mechanical vibrations. Thorne Lay and Terry C. Wallace (1995), following this point of view, define seismology as the study of the generation, propagation, and recording of elastic waves in the Earth (and other celestial bodies), and of the sources that produce them, and conclude that recordings of ground motion as a function of time, or seismograms, provide the basic data for seismologists. A similar approach is followed by Seth Stein and Michael Wyssession (2003). However, Cinna Lomnitz (1994) considers this approach to be rather narrow, because seismograms provide us with much less information about earthquakes than is needed. Moreover, this definition downplays many other aspects present in the complex phenomena of earthquakes.

In a more traditional approach, seismology is defined in a broader sense as the science of the study of earthquakes. The analysis of seismic waves forms a very important part of seismology, but not its totality. Bruce A. Bolt (1978) considers the task of seismologists to be the study of all aspects of earthquakes, including their causes, occurrence, and properties. For Bullen (1947), it is evident that the study of earthquakes belongs to many fields of knowledge, such as physics, chemistry, geology, engineering, and even philosophy. For this

reason many authors consider seismology as a multidisciplinary science (Macelwane and Sohon, 1936; Madariaga and Perrier, 1991).

There is no doubt that the study of seismic waves recorded by seismographs and their physico-mathematical formulation is fundamental to seismology, for example, to the study of the mechanism causing earthquakes and the constitution of the Earth's interior. However, this does not imply that wave analysis is the only source of information about earthquakes. The seismicity of a region, for example, cannot be understood correctly if solely instrumentally recorded earthquakes are considered. Owing to the long return periods for large earthquakes, the study of historical earthquakes is essential. The need to go even farther back into the past has promoted the study of other types of information from archeoseismicity and paleoseismicity. The characteristics of large earthquakes cannot be fully understood without geologic field observations after their occurrence. Comparison of geodesic measurements before and after earthquakes is another important source of knowledge. Modern satellite observations such as from global positioning system (GPS) and interferometric synthetic aperture radar (INSAR) provide very useful seismological information. All these types of data are important in helping one to interpret the nature of earthquakes and their consequences, and must be integrated with information obtained from the analysis of seismic waves.

Two parts of seismology with a marked multidisciplinary character are the evaluation of seismic risk and work toward the prevention and the prediction of earthquakes. In the first case, the interaction of seismologists with geologists and engineers is necessary in order to correctly assess earthquake hazards, expected ground motion, soil conditions, seismic zonation, and the responses of structures and buildings. In the second, many of the suggested precursory phenomena (for example, electromagnetic signals, changes in resistivity, emissions of radon gas, and changes in geodesic measurements) are not directly related to seismic waves. Progress in the difficult problem of earthquake prediction cannot be achieved without a great multidisciplinary effort involving scientists working in many fields, such as seismologists, engineers, geologists, and physicists. In recent years emphasis has been placed on earthquake prevention with the development of detailed seismic risk analyses, robust seismic building codes, and earthquake early warning systems. Depending on correct assessment of the seismic risk and the adequacy of the design and construction of buildings, the damage from earthquakes, especially loss of human lives, may vary greatly. Finally, we must not forget that earthquakes are natural disasters that affect human lives (Zeilinga de Boer and Sanders, 2005; Coen, 2013). The response of the population to the occurrence of an earthquake must also be taken into account, with all its serious psychological, social, and economic consequences. Seismologists cannot be indifferent to all these problems.

1.3 Divisions of seismology

Seismology can be divided into three disciplines: seismology in the strict sense, seismic engineering, and seismic exploration. Seismology treats the occurrence of earthquakes and their related phenomena and is primarily based on application of the principles of the mechanics of a continuous medium, and in particular of the theory of elasticity to them.

As has already been mentioned, its two main subjects are the generation of earthquakes and the vibrations and propagation of seismic waves inside the Earth. From observations of these vibrations, together with other types of data, we derive our knowledge about the nature of earthquakes, the structure of the Earth's interior, and its dynamic characteristics. The part of seismology that deals with seismologic instrumentation, called seismometry, studies the physical theory of the various types of instruments used to measure seismic motion.

Seismic or earthquake engineering is an applied science that deals with how the motion produced by earthquakes affects buildings and other man-made structures. Starting from the characterization of ground displacement, velocity, and acceleration, seismic engineering proceeds to consider their effects on structures and seeks to design them to resist such motions. If earthquake-resistant structures are not to be unnecessarily expensive, a reliable evaluation of the expected ground motion at a particular site is necessary. For this task, an assessment of the seismic risk for a particular zone is needed. This assessment includes consideration of many factors, such as the occurrence of earthquakes near a site, their source mechanism, seismic wave attenuation and soil conditions, and the vulnerability of structures. The complete evaluation of seismic risk implies the statistical analysis of all these factors and requires the collaboration of seismologists, engineers, and geologists.

In seismic exploration, seismological methods are applied to the search for mineral resources, especially oil deposits. These methods are based on the reflection and refraction of artificially generated seismic waves in geologic structures associated with the presence of such deposits. The methods that have proved to be most effective are those based on vertical reflection of waves. Closely spaced distributions of wave generators and detectors together with complex processing of the digital data allow one to obtain detailed images of the upper part of the Earth's crust. The increasing demand for energy resources makes this work more and more important.

1.4 Theory and observations

Just as in all experimental sciences, theoretical and observational aspects of seismology must be considered. The first are based on the principles of the mechanics of continuous media, with the assumption that the Earth is an imperfectly elastic body in which vibrations are produced by earthquakes. The study of the generation of these vibrations constitutes the theory of the source mechanism. In this theory, different models of the processes occurring at the foci of earthquakes are proposed. They range from the more simple instantaneous point sources to the more complex fracture processes. The aim is to approximate the process of fracture and slip that takes place along geologic faults.

Vibrations in the Earth can be treated using two approaches: wave propagation and normal modes theory. The first approach considers waves propagating inside the Earth or on its surface. The second considers the eigenvibrations or oscillations of the Earth as a whole. This second approach is necessary when wave lengths are near the dimensions of the Earth. In the simplest models, the Earth behaves like a homogeneous isotropic perfectly elastic

medium. For some problems the flat-Earth approximation may be sufficient, whereas others require treatment of its sphericity. Heterogeneity in the Earth can be treated using layered models with different elastic properties for each layer, or models in which these properties vary with the spatial coordinates. The assumption of a spherical radially heterogeneous medium is useful in providing a close approximation to the real Earth. Ray theory is employed as a useful high-frequency approximation to wave propagation in heterogeneous media. Surface waves in layered media describe wave dispersion with the separation of phase and group velocities. The lack of perfect elasticity is accounted for by introducing the attenuation of vibrations and waves and by considering viscoelastic models. Isotropic models are adopted as a first approximation but further analysis needs to consider anisotropic conditions. By proceeding through these successive modifications in models of the Earth, its imperfect elasticity, heterogeneity, and anisotropy can be adequately considered.

An important part of seismologic observations consists in the recording of the ground's motion by instruments installed on its surface. Nowadays, classical analog seismograms on photographic paper have largely been replaced by digital data kept on magnetic tapes or compact disks, which can be obtained directly from world data banks through the Internet practically in real time. Previously to their interpretation through the use of digital computers, seismologic observations usually needed careful complex numerical processing. As has already been mentioned, important seismologic data are also provided by other sources, for example, historical records of damage from pre-instrumental earthquakes, field observations of structural damage and ground deformation after earthquakes, geodesic measurements related to the occurrence of earthquakes, *in situ* stress measurements, and geologic and tectonic implications. A modern source of seismological observations is provided by continuous high-rate GPS data of coseismic displacements. GPS records function like a new type of seismogram. Crustal deformations produced by large earthquakes can be obtained by modern INSAR observations. Observations of field geology and damage to structures after earthquakes also provide an important source of information. Progress in methods of observation of all kinds of seismologic data has allowed one to apply models of increasing complexity to the problems of the generation of earthquakes and determination of the structure of the Earth's interior.

The relation between observations and theories or models can be approached through direct and inverse problems. The direct problem refers to the determination of ground displacements from theoretical models of the generation and propagation of seismic waves. In the direct problem, theoretical models are assumed *a priori* and from them synthetic displacements are determined, which are then compared with observations. If they agree, we consider the model well suited to observations. However, in many instances, there is no assurance of its uniqueness and many other models may equally well satisfy the same observations. The inverse problem consists in estimation of the parameters of a theoretical model from observations. This is often a more complicated problem than the direct one. Observations are always incomplete and contain errors, so that a solution of the inverse problem may only exist in a probabilistic sense. In general, inverse problems become more intractable as the number of parameters of the model increases. The mathematics of inverse problems requires, generally, the solution of non-linear integral equations. Linearization of the problem is a standard procedure that leads, very often, to large unstable systems of

equations. Difficulties in the solution of inverse problems lead to their substitution by repeated solutions of direct problems until sufficient agreement between observations and synthetic data predicted by the assumed models is reached.

1.5 International cooperation

The main objectives of seismology require the cooperation of, and exchange of observations among, scientists from different parts of the world. This collaboration was accomplished from early times through private initiatives. The global character of large earthquakes soon required the establishment of institutional cooperation at national and international levels. The first organizations were national ones such as the Seismological Society of Japan, created after the earthquake of 1880 with Milne as first secretary. In 1890, the Committee for the Investigation of Earthquakes was founded, also in Japan, of which Omori was president from 1897 to 1923. In Italy, the Italian Seismological Society (*Società Sismologica Italiana*) was created in 1895; Luigi Palmieri, Timoteo Bertelli, and Giuseppe Mercalli were among its first members. Another national society with great influence in the history of seismology is the Seismological Society of America, which was founded in 1906 as a response to the great San Francisco earthquake, with Charles Davidson as its first president. The idea of an international association of seismology was first proposed by Georg Gerland, during the Sixth International Congress of Geography that was held in London in 1895. In 1904, the International Association of Seismology was finally created, but it was suppressed in 1916. Since 1922, seismology has formed a section of the International Union of Geodesy and Geophysics (IUGG), created in 1919. In 1930, the IUGG was reorganized and included as one of its associations the International Association of Seismology, which finally, in 1951, received its present name of the International Association of Seismology and Physics of the Earth's Interior (IASPEI). One of its commissions is the European Seismological Commission (ESC), which was founded in 1951. There are also active seismology sections of geophysical scientific societies such as the American Geophysical Union (AGU) and European Geosciences Union (EGU).

Exchange of seismologic data between observatories was carried out in the past through the publication of seismologic bulletins. These bulletins preserve a great wealth of information about earthquakes of the early instrumental period. One of the first publications of epicenter determinations was The Reports of the Seismological Committee of the British Association for the Advancement of Science, which started in 1911 with the determinations for the period 1899–1903. In 1922, this publication became the International Seismological Summary (ISS), its first volume being dedicated to the earthquakes of 1918. Later, in 1963, the publication was continued by the International Seismological Centre (ISC), Newbury, UK. The Bureau Central International de Séismologie (BCIS) was created in Strasbourg in 1906 and published a bulletin with epicenter determinations from 1904 until 1975. In 1976, the Centre Séismologique Européen Méditerranéen (CSEM) was created by the ESC with the task of determining hypocenters of earthquakes of the Mediterranean region. Other agencies also started to publish epicenter determinations, such as, in North America, the

Jesuit Seismological Association that was active between 1925 and 1960 and the United States Coast and Geodetic Survey (USCGS), which was later transferred to the National Earthquake Information Center (NEIC), which is dependent on the United States Geological Survey (USGS). Since 1968, its monthly publication Preliminary Determination of Epicenters has included also information on determinations of focal mechanisms for sufficiently large earthquakes. Similar information has also been published since 1977 by Harvard University and was continued from 2006 onwards by the Global Centroid Moment Tensor Project. At present there are several world centers of seismologic data, including digital seismograms from broad-band stations, such as IRIS (USA), GEOFON (Germany), and ORFEUS (Holland).

1.6 Books, journals, and websites

A list of books which cover the different topics relevant in seismology is given in the Bibliography. It is divided into three sections: general seismology, special topics in seismology, and elasticity and wave mechanics. Some of these books are mentioned in the following paragraphs.

Among the early treatises on seismology are those of:

Mallet (1862), *Great Neapolitan Earthquake of 1857: The First Principles of Observational Seismology*.

Rudolf Hoernes (1893), *Erdbebenkunde*.

Milne (1898), *Seismology* (Fig. 1.2).

At the beginning of the last century, several books on seismology were published, among them were those by:

Sieberg (1904), *Handbuch der Erdbebenkunde*.

Hobbs (1907), *Earthquakes. An Introduction to Seismic Geology*.

Montessus de Ballore (1911), *La sismologie moderne*.

Galitzin (1914), *Vorlesungen der Seismometrie*.

From 1930, textbooks about seismology that may be considered modern started to be published. Only those of a general character will be mentioned (full references are given in the Bibliography):

Macelwane and Sohon (1936), *Introduction to Theoretical Seismology*. Part I, Geodynamics and Part II, Seismometry.

Byerly (1942), *Seismology*.

Bullen (1947), *An Introduction to the Theory of Seismology*.

Richter (1958), *Elementary Seismology*.

Sawarensky and Kirnos (1960), *Elemente der Seismologie und Seismometrie*.

Bath (1973), *Introduction to Seismology*.

SEISMOLOGY

BY

JOHN MILNE, F.R.S., F.G.S.

LATE PROFESSOR OF MINING, GEOLOGY, AND SEISMOLOGY IN THE
IMPERIAL COLLEGE OF ENGINEERING, TOKIO, JAPAN
HON. FELLOW KING'S COLLEGE, LONDON

WITH FIFTY-THREE FIGURES

FIRST EDITION

LONDON

KEGAN PAUL, TRENCH, TRÜBNER & CO. LTD.

PATERNOSTER HOUSE CHARING CROSS ROAD

1898

Fig. 1.2

The title page of Milne's book on seismology (1898) (Cambridge University Press).

More recently, since 1979, several textbooks on general seismology at various levels have been published. Four excellent advanced books are:

Pilant (1979), *Elastic Waves in the Earth*.

Aki and Richards (1980), *Quantitative Seismology. Theory and Methods*.

Ben Menahem and Singh (1981), *Seismic Waves and Sources*.

Dahlen and Tromp (1998), *Theoretical Global Seismology*.

At an introductory level there are books by:

Bullen and Bolt (1985), *An Introduction to the Theory of Seismology*.

Bolt (1978), *Earthquakes, a Primer*.

Gubbins (1990), *Seismology and Plate Tectonics*.

Madariaga and Perrier (1991), *Tremblements de terre*.

Lay and Wallace (1995), *Modern Global Seismology*.

Doyle (1995), *Seismology*.

Gershanik (1995), *Sismologia*.

Udías and Mezcua (1996), *Fundamentos de sismología*.

Shearer (1999), *Introduction to Seismology*.

Udías (1999), *Principles of Seismology*.

Pujol (2003), *Elastic Wave Propagation and Generation in Seismology*.

Stein and Wysession (2003), *An Introduction to Seismology, Earthquakes and Earth Structure*.

There are books covering only certain aspects of seismology, such as, for example, ray theory, by Červený (2001), wave propagation and free oscillations, by Officer (1958), Ewing *et al.* (1957), Lapwood and Usami (1981), Kennett (1983), and Babuska and Cara (1991); source mechanisms, by Kasahara (1981), Kostrov and Das (1988), Scholz (1990, 2002), and Udías *et al.* (2014); seismicity, earthquake prediction, and other topics such as extraterrestrial seismology, by Gutenberg and Richter (1954), Kisslinger and Zuzuki (1977), Kulhanek (1990), Lomnitz (1994), and Tong and García (2015).

There are excellent collections of review papers, such as those by Dziewonski and Boschi (eds.) (1980), Kanamori and Boschi (eds.) (1983), Boschi *et al.* (eds.) (1996), and Kanamori (ed.) (2009). Entries on seismologic subjects in James (ed.) (1989) *The Encyclopedia of Solid Earth Geophysics* are very good short up-to-date presentations.

The first scientific articles about seismology were published in the *Bollettino del vulcanismo Italiano* founded by de Rossi in 1874 and in the *Beiträge zur Geophysik*, founded by Gerland in 1887. The first journals exclusively dedicated to seismology were the *Transactions of the Seismological Society of Japan*, published from 1880 to 1892 and the *Seismological Journal of Japan*, published from 1892 to 1895, both directed by Milne. In 1985, the *Bollettino della Società Sismologica Italiana* was founded by the Italian Seismologic Society, and, in 1897, the *Mitteilungen der Erdbeben* was founded by the Vienna Academy of Sciences. In 1907, publication began of the *Bulletin of the Imperial Earthquake Investigation Committee* in Japan, in 1908, the *Publications du Bureau Central de l'Association Internationale de Sismologie* started publishing in Strasbourg, in 1911, the *Bulletin of the Seismological Society of America*, in 1926, the *Bulletin of the Earthquake Research Institute* of Tokyo University, and, in 1929, *Earthquakes Notes*, which in 1987 changed its name to *Seismological Research Letters*. In 1997, publication began of the *Journal of Seismology* (Springer, Dordrecht). Some examples of journals dedicated to the field of earthquake engineering are: *Earthquake Engineering and Spectral Dynamics*

(1997), *Soil Dynamics and Earthquake Engineering* (1981), *Earthquake Spectra* (1985), *Natural Hazards* (1988), and *Bulletin of Earthquake Engineering* (2003).

Besides journals dedicated entirely to seismology, articles on this subject are published in geophysical journals. The list is very long so only the most representative are mentioned, in chronological order of the first year of publication: *Geophysical Magazine* (1926), *Fizica ziemly* (1937), *Pure and Applied Geophysics* (*Geofisica pura e applicata*) (1939), *Annali di geofisica* (1948), *Journal of Geophysical Research* (1949), *Journal of Physics of the Earth* (1952), *Advances in Geophysics* (1952), *Geophysical Journal International* (1992) (a fusion of the *Geophysical Journal of the Royal Astronomical Society* (1958), the *Zeitschrift für Geophysik* (1924), and the *Annales de Géophysique* (1948)), *Reviews of Geophysics* (1963), *Tectonophysics* (1964), *Earth and Planetary Science Letters* (1966), and *Physics of the Earth and Planetary Interiors* (1967).

Actually, the amount of published material in seismology continues to increase considerably. Students will find it useful to consult the most important textbooks given in the Bibliography, where they will find different approaches to topics treated in this book. Also, they should read some of the classical papers, references to which are given in the References list, and look through the recent issues of seismologic journals to find out about present topics of research.

Information about earthquakes may be found at several websites. This information may be general information such as focal parameters (date, origin time, hypocenter, and magnitude) or more specific, such as wave forms, instrumental responses, studies of some large earthquakes, etc. Some examples of these websites are:

www.usgs.gov/ (U.S. Geological Survey, USA); <http://earthquake.usgs.gov/contactus/golden/neic.php> (National Earthquake Information Center, NEIC, USA); www.iris.edu/hq/ (Incorporated Research Institutions for Seismology, IRIS, USA); www.emsc-csem.org/#2 (EMSC/CSEM, European Mediterranean Seismological Centre, EU); www.orfeus-eu.org/ (Observatories and Research Facilities for European Seismology, ORFEUS, EU); <http://geofon.gfz-potsdam.de/> (Geoforschungszentrum, GEOFON, Potsdam, Germany); <http://geoscope.ipgp.fr/index.php/en/> (GEOSCOPE, Institut de Physique du Globe, Paris, France); www.jma.go.jp/en/quake/ (Japan Meteorological Agency, Earthquake Information, Tokyo, Japan); www.eri.u-tokyo.ac.jp/en/ (Earthquake Research Institute, University of Tokyo, Japan); www.globalcmt.org/CMTsearch.html (Global Centroid Moment Tensor, GCMT, Harvard, USA).

1.7 Summary

Seismology, the science of earthquakes, has a long history as since antiquity man has tried to describe the effects of these natural disasters and study their causes. Modern treatises on seismology began towards the middle of the nineteenth century with the application of the principles of mechanics to the shaking of the Earth produced by earthquakes and the development of seismographs to measure and record this, making possible its quantitative

study. Since the middle of the twentieth century seismology has experienced a rapid development in its theoretical and applied aspects.

Seismology covers the study of theoretical and observational aspects of earthquakes. These are related to the generation of earthquakes through study of their source mechanism and the propagation of seismic waves through the Earth that provides knowledge about its structure and dynamics. Practical applications of seismology include the important subject of assessment of seismic risk. Seismology must be considered to be a multidisciplinary science that includes all aspects related to the occurrence of earthquakes and not only those of the generation and propagation of seismic waves. Thus, contributions from, for example, geology, geodesy, and engineering must also be considered and the human aspects related to these natural disasters must not be forgotten.

An important aspect of seismology that is mentioned in this introductory chapter is the international cooperation necessary for the study of earthquakes and the establishment of scientific organizations of international and national character. Finally, a list is given of some of the most important textbooks, specialized journals, and websites where information can be found.

The first approach to the study of earthquake occurrence is from a descriptive point of view. Following this approach, we present in this chapter a first view of the basic ideas about earthquake occurrence and damage, seismic intensity as a measure of the size of earthquakes, earthquakes in relation to geological faults, their geographical and time distribution, some basic ideas of the Earth's structure and dynamics related to earthquake occurrence, and of seismic hazard and risk. This short presentation serves as an introduction to and shows the need for a physical mathematical approach to the quantified study of earthquakes, which is the subject of subsequent chapters.

2.1 Earthquakes: natural disasters

Earthquakes, in some cases followed by *tsunamis* (Japanese for bay waves), landslides, and fires, are, first of all, natural disasters that take the greatest number of human lives and produce the heaviest material damage ([Appendix 7](#)). Reports about large earthquakes and their catastrophic consequences have come to us since ancient times. Historians have left us vivid descriptions of the occurrence of large earthquakes throughout human history. Some of the oldest earthquakes for which we have historical evidence are from as far back as a thousand years BC. Some large earthquakes have had important consequences in human history; for example, an earthquake in 464 BC ended the hegemony of Sparta over Greece ([Zeilinga de Boer and Sanders, 2005](#)). The great Lisbon earthquake of 1755, the largest ever felt in Europe, which was followed by a large tsunami and caused some 15,000 casualties, gave origin to an intense philosophical and theological debate and is considered to mark the origin of seismology as a science, as we have seen in [Chapter 1](#). Disasters caused by earthquakes are a continuous reminder of the vulnerability of human constructions. Even over the last ten years there have been a considerable number of very damaging large earthquakes with extremely high loss of human life. The Sumatra earthquake of 2004 produced about 228,000 deaths, most of these in the following large tsunami that struck the coasts of the Indian Ocean. In 2010 an earthquake destroyed Port-au-Prince, the capital of Haiti, causing some 316,000 deaths due to the low quality of its buildings ([Fig. 2.1](#)). Reconstruction after five years is only progressing at a very slow pace. In spite of preparation against earthquakes in Japan, an unexpected tsunami with waves of over 20 m following the large offshore Tohoku earthquake of 2011 caused 28,050 deaths. The damage produced at the Fukushima nuclear power plant added serious dangers and concerns about the safety of this type of power plant. The recent Nepal earthquake of 2015 with more than

**Fig. 2.1**

Damage of the 10 January 2010 earthquake $M_w = 7$, in Port-au-Prince, Haiti (U.S. Geological Survey). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

9000 deaths caused heavy destruction in the capital Kathmandu and left some villages buried by landslides. These chilling examples of the consequences of some large earthquakes within the past ten years highlights the importance of these phenomena and of their study ([Appendix 7](#)).

In these examples of recent large earthquakes, damage and casualties have been caused in addition to the shaking of the ground by tsunamis and landslides. Shaking of the ground, the primary effect of earthquakes, that may reach horizontal accelerations larger than that of gravity, causes total ruin of poor quality buildings, as was the case in Haiti and Nepal, burying many people under their ruins. Seismic resisting construction considered in the building codes used in developed countries can diminish damage and avoid the total collapse of buildings saving most lives, but construction in poor countries remains very vulnerable. Tsunamis are waves produced by the sudden displacement of the ocean bottom by offshore earthquakes which may reach heights of up to 30 m. They may propagate across the ocean for long distances producing damage far away from the earthquake area. In the Pacific Ocean tsunami alert systems exist that permit warnings of their anticipated arrival which allows the evacuation of coastal areas. Only since the Sumatra earthquake have these also been installed in the Indian Ocean. Other effects of earthquakes are large landslides, for example that produced by the 1970 Peru earthquake. A large mass of mud and snow from the Huascaran mountain completely buried the town of Yungay, killing and burying as many as 20,000 people. Damage produced by landslides is difficult to prevent other than by prohibiting building near unstable mountain slopes, but this is rarely possible to achieve. Fires are an added factor that increase the damage caused by earthquakes. A tragic example

is the great Kanto earthquake of 1923, where violent and widespread fires in Tokyo and Yokohama caused most of the 142,800 deaths. Fire also broke out after the renowned San Francisco earthquake of 1906, which lasted for three days and produced greater damage than the shaking itself. Therefore a number of factors besides shaking of the ground in earthquakes contribute to damage and casualties. The number of deaths in an earthquake is a result not only of its size, but also of many other factors. An earthquake with one of the reported highest numbers of casualties was that of Tangshan (China) in 1976, with estimates of numbers of deaths ranging between 250,000 and 650,000, but it was not of the largest magnitude. Estimates of around a total of two million deaths due to earthquakes have been made for the twentieth century. Earthquake losses are likely to increase in the future due to population growth in high-risk seismic areas and to inadequate building quality in low and medium development regions. These considerations show the importance of the study of earthquakes and the need for work to mitigate their effects.

Earthquakes are also very frightening human experiences. They shake the ground that we normally consider to be solid and stable. We therefore feel that if the stable ground fails, then everything else fails with it and we must run to look for refuge as all is being shaken around us. As Charles Darwin wrote after the 1835 Chile earthquake: “A bad earthquake at once destroys the oldest associations; the world, the very emblem of all that is solid had moved beneath our feet like a crust over a fluid.” Lomnitz (1994) talking about earthquakes comments: “Disasters cut deeply into our experience of human life, they bring out the best in humans as well as the worst.” Thus, the image of an earthquake is also used as a metaphor for some events which shake the ground of our position in society, religion, philosophy, or politics. These types of “earthquake” happen in many aspects of our lives, but the real thing with its terrible consequences is at the base of the metaphor. The Bible, for example, refers to the occurrence of large earthquakes to express the great disasters before the end of time: “And there were flashes of lightning, rumblings, peals of thunder, and a great earthquake such as there had never been since man was on the Earth, so great was that earthquake.” (Revelation 16:18). Poets sometimes look for ways and images to express the impression produced by earthquakes. Pablo Neruda describes the large Chilean earthquake of 1960 as an irascible horse kicking the land. Earthquakes will naturally continue to occur, but we have the obligation to contribute to mitigate their catastrophic effects and to ensure public safety. A new concept concerning the response to earthquakes is *seismic resilience*, which includes protection of the lives of citizens during earthquakes, improvement of the capacity of society to respond to earthquakes, preparation to recover quickly after an earthquake, and protection of the economy (Jones, 2015).

Seismology, the science of earthquakes, tries to understand the nature of these phenomena using a variety of approaches. In quantitative seismology, the subject of this book, we apply the principles of physics, especially of the mechanics of elastic bodies, to the motion produced by earthquakes. Quantitative observations of this motion are provided by seismological and other types of instruments (Chapter 3). Application of basic principles of mechanics to motion produced by earthquakes is presented in Chapter 4. Propagation of seismic waves and Earth free oscillations are studied in Chapters 5 to 10 (body waves), Chapter 12 (surface waves), and Chapter 14 (free oscillations). Effects due to anelasticity and anisotropy are studied in Chapter 15. Moreover, analysis of the seismic waves produced

by earthquakes and propagated through the Earth also provides crucial knowledge about the structure of the Earth's interior and its dynamics ([Chapters 11 and 13](#)).

2.2 Size of earthquakes

Earthquakes occur in various sizes, from small tremors only detectable by instruments to large catastrophic events that affect large areas and cause extensive damage and many casualties. While only one or two very large earthquakes may happen worldwide in a year, there are about 200 of moderate size and more than 10,000 light ones. In terms of the energy released at the focus, a very large earthquake corresponds to about 10^{18} J, while a perceptible small one releases only about 10^8 J. One large earthquake releases the energy of about 1000 nuclear explosions of one megaton. This is approximately equivalent to the total global energy consumption in one year. This shows the enormous energy released in large earthquakes. Small earthquakes may only be equivalent to around a few hundred kilograms of TNT, so there is a large difference between large and small earthquakes.

The first way to describe the size of an earthquake is in terms of the form by which it is felt by people and the observations of damage caused to buildings and other structures and of the ground effects, such as fractures, cracks, and landslides. The first attempt to classify the size of earthquakes by the damage caused was that of Domenico Pignataro in Italy in 1783, who described them as slight, moderate, strong, and very strong. A further development was the definition of *seismic intensity* used to classify earthquakes by their size, also based on these kinds of observations. Traditionally, intensity is represented by degrees given by Roman numerals, using scales with generally 10 or 12 degrees (I to X or XII) in which each degree is defined in a descriptive way. The first widely accepted scale of intensity was developed by de Rossi and François-Alphonse Forel in 1883, with ten degrees. A modification of this scale was proposed by Mercalli, still with ten degrees, extended later by Adolfo Cancani and Sieberg to twelve degrees. Although intensity applies directly to the degree to which an earthquake is felt or damage produced at a particular location, it can also be used to designate the size of an earthquake. For this purpose the *maximum intensity* I_{max} or *epicentral intensity* I_0 is used. These two concepts are generally considered to be equivalent, but this need not always be correct. For offshore earthquakes the maximum damage is on the coast, so I_{max} does not correspond to I_0 . For example, the 1985 Mexico earthquake with the greatest damage (I_{max}) in Mexico City had an offshore epicenter (I_0) 325 km away. The location of the instrumental epicenter, in some cases, does not correspond to the region of maximum damage or maximum felt intensity and hence these two values (I_0 and I_{max}) are different. Maximum felt intensity was the first estimation used to classify the size of an earthquake. Even today maximum intensity is used to describe the size of historical earthquakes for which there are no instrumental data. A description of modern day intensity scales is given in [Chapter 16](#).

The development of seismographs to measure ground displacement produced by seismic waves generated by earthquakes led to a classification of the size of earthquakes based on this type of measurement. In 1935, Charles F. Richter developed the concept of *magnitude*

of an earthquake, based on instrumentally measured amplitudes of recorded seismic waves. This is the first quantification of the size of earthquakes. Magnitude is related to the logarithm of the energy released at the focus of an earthquake. According to the different types of scales of magnitude, very large earthquakes have magnitudes 8 or 9 (for example, the 2011 Tohoku, Japan, earthquake, $M_w = 9.0$). The definition of the different scales of magnitude and their relations with intensity and energy will be explained in [Chapter 16](#). A more modern concept of the size of an earthquake is the *seismic moment* defined in 1966 by Keiiti Aki. The seismic moment is related to the area and displacement of the fault that breaks in an earthquake and will be explained in [Chapter 16](#).

2.3 Earthquakes and faults

The occurrence of earthquakes is connected with tectonic processes in the Earth's crust which often result in dislocations and fractures. Systematic studies of field observations after the occurrence of earthquakes in the nineteenth century, such as those of George K. Gilbert (Owen, California, 1872), Bunjiro Koto (Nobi, Japan, 1891), and Richard D. Oldham (Assan, India, 1897), began relating earthquakes with motion along faults in the Earth's crust. With the increase in numbers of field observations and of precision in localization of the epicenters, the relation between earthquakes and faults became more evident. Thus, the cause of earthquakes began to be assigned to the release, by brittle fracture on faults, of the tectonic stresses accumulated in the Earth's crust.

Thus, it can be stated that most earthquakes are of tectonic origin and take place on faults. Determination of fault parameters from seismic observations will be seen in [Chapter 17](#). Non-tectonic earthquakes such as those of volcanic origin or induced by dams have different characteristics ([Chapter 18](#)). Geological field observations show the presence of many fault traces exposed at the Earth's surface, with lengths which vary from some meters to many kilometers ([Fig. 2.2](#)). In general, observed fault traces are not continuous straight lines, but they have bends, bifurcations, offsets, and other complexities. However, a useful simplification is to consider that earthquakes happen by ruptures of various sizes on fault planes. Very large earthquakes break long lengths, for example the Sumatra earthquake of 2004 broke a long fault of about 1000 km with relative displacements of about 10 m. Small earthquakes may break only a few meters with displacements of centimeters.

The depth extent at which faults break by brittle fracture generating earthquakes can be deduced from the depth of shallow earthquakes. From this evidence it has been shown that brittle fracture extends only to around between 15 km and 30 km in depth. This shows that only the upper part of the crust has a brittle nature and can, in consequence, be fractured. This brittle upper part of the crust is called the *seismogenic layer*, or layer capable of generating earthquakes. The lower part of the crust, owing to the increase of temperature and pressure, behaves as a plastic material, which can be deformed without breaking. This behavior of the lower crust has been confirmed from deep seismic reflection studies. This limitation of the depth imposes a condition on the shape of the fault surface, given by the *aspect ratio* L/W (length over width). For small earthquakes with fracture lengths less



Fig. 2.2 Surface rupture of the San Andreas fault, California. View along Carrizo Plains (U.S. Geological Survey, photo R. E. Wallace). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

than 20 km, there is no limitation to the propagation of fracture in all directions, so the shape will be symmetrical, approximately circular or square. For large earthquakes the length is longer than the limit of depth and the shape will be rectangular or elliptical, with longer length than width ([Chapter 19](#)). Very large crustal earthquakes have lengths extending to hundreds of kilometers and only a limited width, less than about 30 km. However, large shallow earthquakes in *subduction zones* (see plate tectonics) with lengths longer than 400 km may have larger widths of around 100 km. The shape of the fault depends, then, on the size and type of earthquake with different shapes for small and large events. A different problem is presented by the nature of earthquakes with foci at depths below 30–60 km (the range varies according to the crust thickness). These earthquakes are produced in regions where tectonic processes have introduced crustal

material inside the Earth's mantle. Intermediate depth ($30\text{--}60\text{ km} < h < 300\text{ km}$) and deep ($300\text{ km} < h < 700\text{ km}$) earthquakes cannot be produced by brittle fracture, owing to the high temperature and pressure existing at those depths. Different processes, like sudden phase transitions, have been proposed for these, though they do happen on weakness surfaces similar to those of faults (Chapters 19 and 21).

In general, earthquakes occur on pre-existing fault planes, although, naturally, at some time faults have to be newly broken. The material between the two moving surfaces of a fault is usually filled with crushed and highly deformed rocks, called *fault gouge*, in a zone that may be up to several meters wide in some places. This material is formed by the repeated differential movement of the two sides of the fault over a long time, which breaks and shears the rock into a fine granular and powdery form, which can be altered by the addition of water into clays and sandy silts. Earthquakes do not occur at present on all observed geological faults, that is, not all faults observed today in the field are *seismically active*. On a given active fault, earthquakes of different size break partial sections of its surface. On a long fault several large earthquakes may be needed to break its whole extent. Faults can also move slowly without producing an earthquake; this type of motion is called *fault creep*.

2.4 Spatial distribution of earthquakes

The distribution of earthquakes and their characteristics within a particular region is known by the term *seismicity*, probably used for the first time by Montessus de Ballore in 1906. The most important aspects of seismicity are given by the geographic distribution of the location of the foci of earthquakes, their size, their occurrence over time, their mechanisms, and the damage produced by them. Studies of seismicity are, then, based on seismic catalogs that include earthquake parameters such as the dates and times of occurrence, focal coordinates, size, given by magnitude or maximum intensity, and other characteristics such as focal mechanisms and their correlation to regional geological and geophysical features. Historical descriptions of earthquakes can be traced back to written records of old civilizations such as those of China and India and, for Europe, of Greece and Rome. Among the first universal catalogs of earthquakes are those published by Bonito in 1691, Zahn in 1696, and Moreira de Mendonça in 1758. Modern world catalogs began with the work of Karl Adolf von Hoff in 1840, Robert and John W. Mallet in 1858, and Milne in 1911. Today, global and regional catalogs are being compiled by various agencies, such as the ISC (International Seismological Centre, Newbury, UK) and NEIC (National Earthquake Information Service, Denver, USA) (see Chapter 1). Among the first studies of seismicity for some regions are those by Alexis Perrey and Montessus de Ballore between 1850 and 1923, Sieberg between 1923 and 1933, global seismicity by Gutenberg and Richter (1954), and European by Karnik (1969, 1971). Studies of seismicity are fundamental for understanding the relation between earthquakes and tectonics, the geodynamic conditions of a region, and for assessment of the danger of the occurrence of earthquakes or *seismic hazard* and *risk* (Chapter 21).

Since antiquity it has been known that some regions are more prone to the occurrence of earthquakes than others. Thus we can separate seismically active regions from those that are more stable. Once the determination of the geographical location of earthquakes had become sufficiently accurate, it was observed that *epicenters* of earthquakes (the epicenter is the horizontal projection on the Earth's surface of the focus of an earthquake) occur in narrow zones that surround relatively stable regions (Fig. 2.3). These bands or alignments of earthquakes coincide in some cases with the margins of continents, or with oceanic ridges under the oceans. However, not all continental margins correspond to seismic zones, so that they can be divided into *active* and *passive margins*. This classification is related to *plate tectonics*, as we will see later. For example, the western margin of the Americas is active whereas the eastern margin is passive, the northern margin of Eurasia is passive whereas its southern and eastern margins are active, and all margins of Africa are passive.

The depths of earthquakes, as already mentioned, vary from the surface to about 700 km. According to their depth, earthquakes are classified as shallow (less than 30–60 km), that is, inside the crust, intermediate (30–60 to 300 km), and deep (more than 300 km). As a general rule, in oceanic ridges there are only shallow earthquakes whereas in island arc regions there are shallow, intermediate, and deep earthquakes (see the islands arcs of the Pacific Ocean and the ridges in the Atlantic and Pacific in Figs. 2.3 and 2.4). Although shallow earthquakes are classified as those less than 30–60 km deep, most of them take place at depths less than 20 km, that is, in the upper rigid part of the crust (Chapter 16). Earthquakes at greater depths correspond to zones where this rigid and relatively cold material of the upper crust is introduced into the upper mantle.

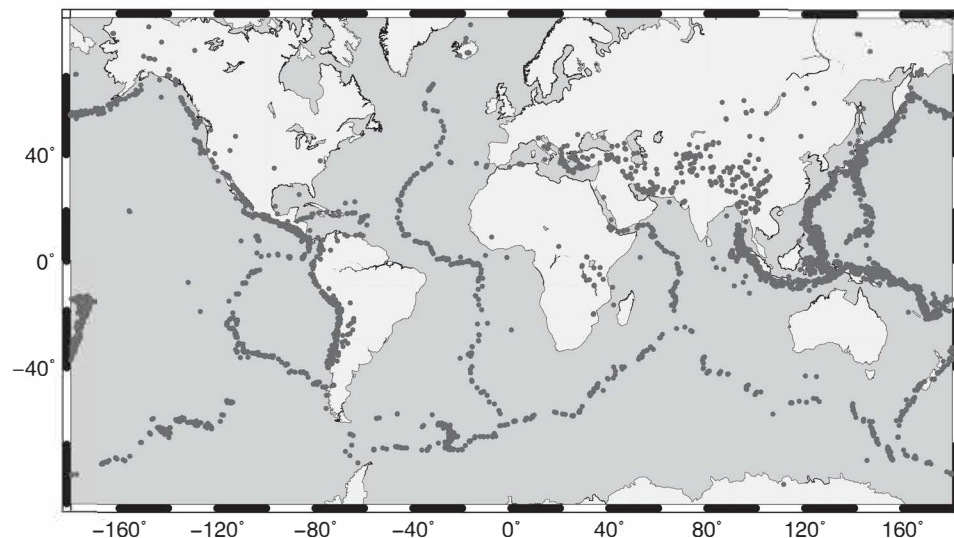


Fig. 2.3

World map of shallow earthquakes ($h < 60$ km) for the period 2000–2015, $M > 5$ (data from NEIC, U.S. Geological Survey). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

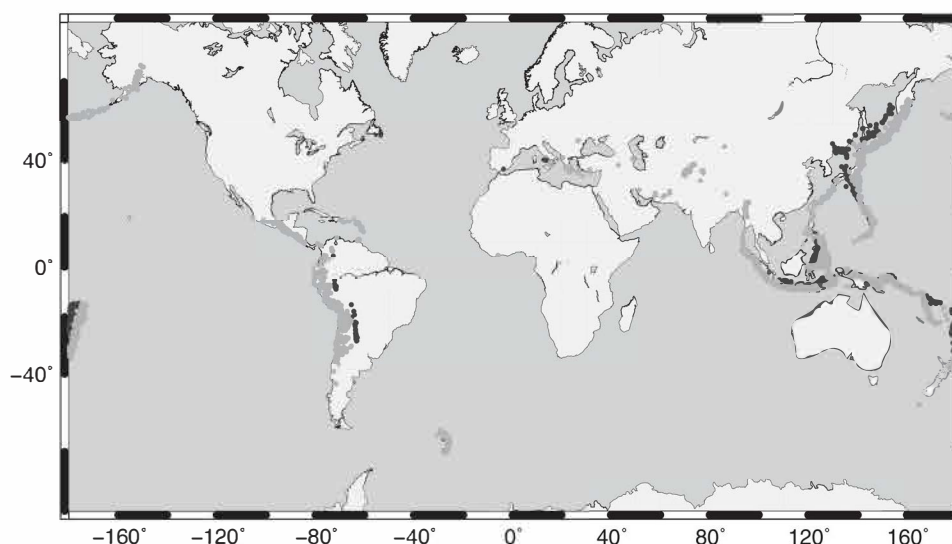


Fig. 2.4

World map of intermediate depth ($60 \text{ km} < h < 300 \text{ km}$, green) and deep ($300 < h < 700 \text{ km}$, blue), earthquakes period 2000–2015, $M > 5$ (data from NEIC, U.S. Geological Survey). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

2.5 Temporal distribution of earthquakes

In studies of seismicity, the temporal distribution of earthquakes is as important as their spatial distribution. In a very general way, it can be said that earthquakes in a region behave as a temporal series of point events resulting from a process of release of stress in the Earth's crust. Statistical studies of temporal series of earthquakes reveal some characteristics of their distribution. The limits of the area being studied and the duration of time intervals are two important factors to consider.

From a statistical point of view, the simplest model of the occurrence of earthquakes with time is a *Poisson distribution*. This distribution assumes that earthquakes are independent events, that is, their occurrence does not affect that of others. Studies of temporal series of earthquakes show that they deviate, to greater or lesser extents, from Poisson's law. The global distribution of large earthquakes follows this distribution in time quite well, but moderate and small earthquakes and those in a limited region do not. The assumption of independence is, then, not totally correct since earthquakes are grouped in series of various kinds with interrelated shocks. The interdependence of earthquakes shows up in the occurrence of clusters such as sequences of aftershocks and swarms. Clusters can be better observed in temporal series including small earthquakes in limited areas. Shocks that are near in space and time are necessarily interrelated.

The most important clustering of earthquakes is that associated with the occurrence of an earthquake of large magnitude which is called the *main shock*, and may sometimes be preceded by some shocks and generally followed by a series of smaller ones. Earthquakes

of lesser magnitude preceding just the main shock are called *foreshocks* and those following immediately after are called *aftershocks*. The occurrence of aftershocks is a common phenomenon and is related to the release of energy in the fracture zone that is not completed by the main shock. At the source volume not all of the accumulated stress is released by the main shock, but rather zones of the source remain unbroken and break afterwards, causing aftershocks. Foreshocks are less frequent, but many earthquakes are preceded by small shocks that break weak zones on the fault plane before the main event. When, in a series of earthquakes in a small area, there is no a main shock, the series is called a *swarm of earthquakes* (Chapter 21).

In a general way, series of earthquakes can be related to the nature of the conditions of the material of the fracture zone (Mogi, 1963). If the material is very homogeneous and the distribution of stress uniform, there are no foreshocks and the main shock is followed by a series of aftershocks of lesser magnitude. Foreshocks are associated with heterogeneities in the source material that result also in longer sequences of aftershocks. If the material is very heterogeneous and the stress not uniform, then earthquakes happen in swarms without a real main shock.

Both the spatial and the temporal distribution of earthquakes within series of aftershocks are of interest. Aftershocks are distributed over the fracture area of the main shock and, in some cases, there is a certain migration from one end of the fault to the other. Thus, the fault area of large earthquakes can be determined by the area over which the aftershocks are distributed. The number of aftershocks, their magnitude, and their duration depend on the magnitude of the main shock. An earthquake of moderate size (magnitude 5) usually has a sequence of aftershocks lasting a few days, whereas for a large one (magnitude 8) the series may last more than a year. The 2004 Sumatra–Andaman earthquake ($M = 9$) had a long series of aftershocks lasting more than a year and a half, over an area about 1500 km in length (Fig. 2.5). To estimate the total duration of a series of aftershocks is sometimes difficult, especially for very large earthquakes. The last part of the series may last for years and it is difficult to establish when the seismicity recovers to its normal level. In general, shallow shocks have longer series of aftershocks than do deeper ones. The magnitudes of aftershocks depend on that of the main shock. The largest aftershock in a series has a magnitude about one (or 1.2) unit lower than that of the main shock (this is known as Markus Båth's law). Another interesting phenomenon in the occurrence of earthquakes is that of *earthquake doublets*, when the main shock is followed after a short time interval by another shock of similar magnitude located very near to it. The term earthquake doublet is not clearly defined and different mechanisms have been proposed for its occurrence. These problems will be seen in more detail in Chapter 21.

2.6 Earth's structure: crust, mantle, and core

The occurrence of earthquakes produces elastic waves that propagate through the interior of the Earth. Analysis of these waves provides valuable information about the structure of the Earth. One can say that earthquakes with their waves illuminate the interior of the Earth and allow us to look into its structure. This application of the knowledge provided by

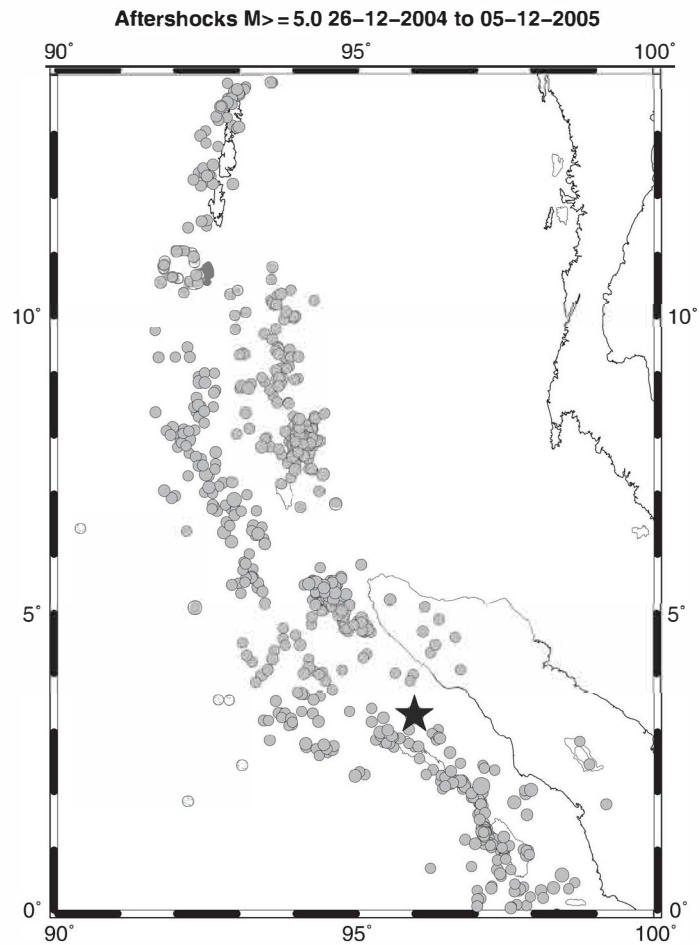


Fig. 2.5

Aftershocks with $M > 5$ during the year after the Sumatra earthquake of 26 December 2004, $M_w = 9.1$. The star shows the location of the main shock (data from NEIC, U.S. Geological Survey).

earthquakes constitutes a very important part of seismology. As we will see, two types of waves are produced by earthquakes: body waves which propagate through the interior of the Earth and surface waves on the surface. Body waves are of two types: compressional or P waves and shear or S waves. The theory of the propagation of body waves in various types of media will be seen in [Chapters 5 to 11](#) and surface waves will be treated in [Chapters 12 and 13](#). Studies using different types of seismic waves have provided us with detailed knowledge of the structure of the Earth from its surface to its center. At present, the most complete models of the Earth are those with non-radial symmetry (lateral heterogeneities), anisotropy, and anelasticity. Details of these studies are shown in [Chapters 11, 13, and 14](#).

The outermost part of the Earth is called the *crust*. From a seismological point of view, the first observations that led to the discovery of the existence of a sharp contrast between the material of the crust and that below it, called the *mantle*, were made in 1909 by Andrija Mohorovičić, who studied travel times of earthquakes in central

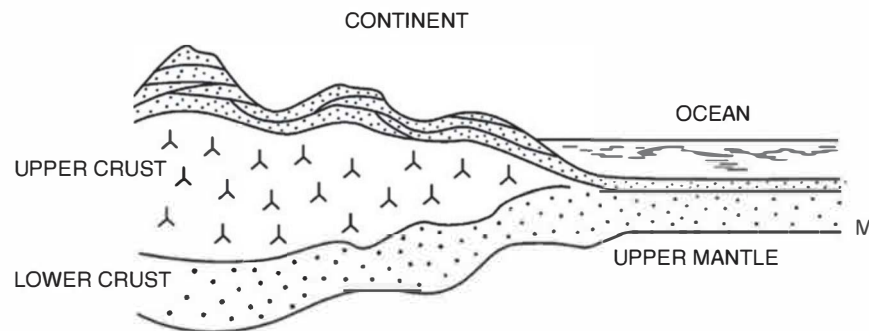


Fig. 2.6

Variation of the crust's thickness on going from oceanic to continental structure.

Europe. He observed that they show the presence of a velocity discontinuity at a depth of about 30 km, which he identified as the base of the crust. This discontinuity today bears his name and is called, in abbreviated form, the *Moho*. A second discontinuity inside the crust was discovered by Victor Conrad in 1923 (it is called the Conrad discontinuity) and Jeffreys in 1926. The thickness of the crust is not homogeneous; in shields or stable continental zones it is about 30 km, whereas under the oceans it is only 8–15 km, and in regions of high mountains it can be as much as 60 km (Fig. 2.6). The simplest models of the continental crust have two layers with constant velocity or with a small gradient, covered by a thin layer of sediments. These two layers were known as granitic and basaltic layers due to their assumed compositions and today are more generally referred to as the upper and lower crust.

The material under the crust down to a depth of 2900 km forms the *mantle*. The mantle is divided into two regions, the *upper* and *lower* mantle. The upper mantle extends from under the crust to a depth of 700 km. Velocities of waves increase with depth from the values under the Moho. Approximately between 60 and 220 km depth, a low-velocity layer, which is more pronounced for S waves, is found. The depth, thickness, and decrease in velocity of this layer vary from region to region and in some regions it may even not be present. Under the low-velocity layer, from approximately 220 km depth, velocities both for P and for S waves increase slowly, with two zones of rapid increase in velocity at depths of 450 and 670 km. The lower mantle extends in depth from 700 km to the boundary of the *core* at 2900 km (the core–mantle boundary, CMB, Fig. 2.7). The seismic characteristics of the lower mantle are very uniform, with a gradual increase in P and S wave velocities. Details about the composition of the mantle will be seen in Chapter 11.

The first evidence of the existence of the Earth's core from the analysis of traveling times of seismic waves was presented in 1906 by Oldham, who deduced that its velocity is lower than that of the mantle. In 1912, Gutenberg fixed the depth of the core at 2900 km, from the study of reflected waves. Transverse waves do not propagate in the outer core, which indicates that its material is fluid enough not to allow the existence of shear stress, as shown by Jeffreys in 1926. The existence of a solid *inner core* was first proposed in 1936 by Inge Lehmann. The Earth's core is, then, formed by two regions, an *outer core* of

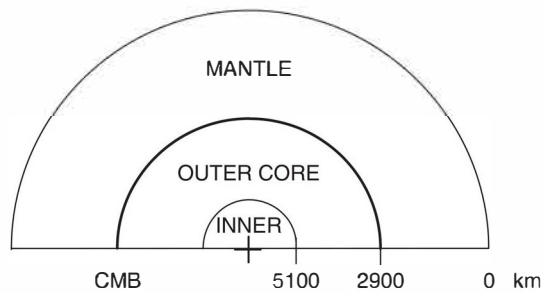


Fig. 2.7 Main regions of the interior of the Earth under the crust (CMB, core–mantle boundary).

3486 km radius and an inner core of 1216 km radius (Fig. 2.7). Details of the structure and wave propagation in the core will be shown in Chapter 11.

2.7 Plate tectonics

The relation between geological characteristics and the occurrence of earthquakes began to be studied at the end of the nineteenth century and very early on in the twentieth century, by Milne, Mercalli, Sieberg, Montessus de Ballore, Koto, and Omori, among others. The term *seismotectonics* started to be used in around 1910 by Sieberg and Montessus de Ballore and is applied to the characteristics of the occurrence of earthquakes in relation to regional tectonics and general geodynamic conditions. In seismotectonic studies one tries to integrate earthquake data with other information available from the tectonics, geophysics, and geology of a particular region. Seismologic data used in seismotectonics include the geographic distribution of epicenters, with indications of their magnitude, depth, and focal mechanism. The tectonic interpretations of earthquakes, included in what are called seismotectonic studies, depend on accepted theories for the general processes active in the Earth's crust. Historically, the first of these theories, presented by James Dana, James Hall, and Suess, among others, were based on vertical movements of crustal blocks related to the contraction of the Earth due to its cooling. Authors of early seismotectonic studies used this tectonic framework to interpret seismic data. In 1912, Alfred Wegener proposed his theory of *continental drift*, whereby tectonic processes are related to the horizontal motion of continental blocks, but the geophysical objections presented against this theory (Jeffreys, 1959) were an obstacle to its use in the interpretation of earthquake data. During the 1960s, *plate tectonics theory*, in whose development seismologic data played an important role, was introduced and, ever since, modern seismotectonic studies have been based on its principles.

The basic ideas of plate tectonics can be summarized in a very brief form as follows (see for example Kearey and Vine, 1996). First of all, the basic unit of tectonic motion is the *lithosphere–asthenosphere* system. The lithosphere is a layer of thickness about 100 km that includes the crust and part of the upper mantle which behaves as a rigid and cool

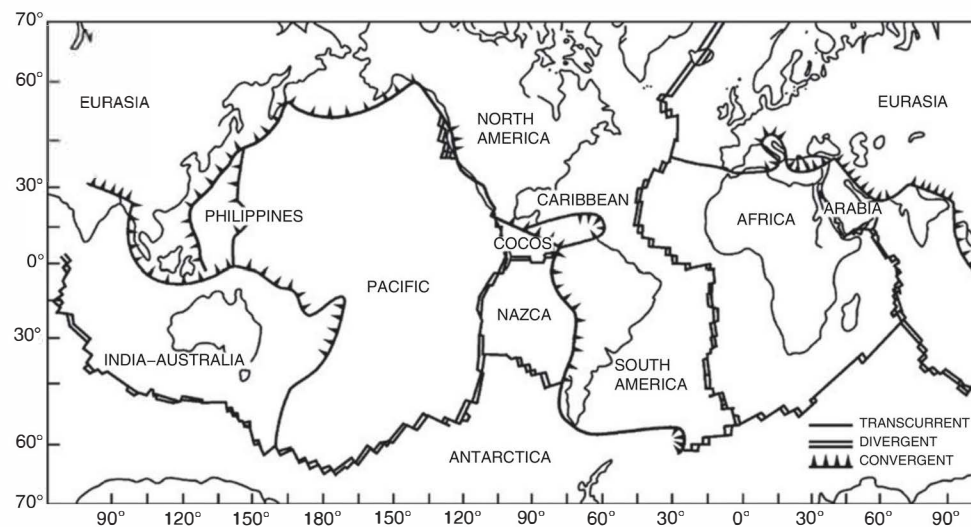


Fig. 2.8 Principal lithospheric plates and types of plate boundaries.

material in relation to the underlying layer. The asthenosphere is a weak, warmer layer that behaves with plastic or viscous flow and corresponds to the low-velocity zone of seismic waves. The viscosity of the asthenosphere allows horizontal motion of the lithosphere, with velocities in the range $1\text{--}8\text{ cm yr}^{-1}$. The motion of the lithosphere is due to a number of processes, including thermal convection currents inside the mantle. The lithosphere is divided into a number of plates that, generally, include both continental and oceanic crust (Fig. 2.8). The most important plates are the Pacific, the Americas (sometimes divided into two, North and South), Eurasia, Australia–India (sometimes also divided), Africa, and Antarctica. Other minor plates are Nazca, Cocos, Philippines, Caribbean, Arabia, Somalia, and Juan de Fuca. Other even smaller units are often called sub-plates or microplates. Plate boundaries can be established from the distribution of seismic regions, as was mentioned in the discussion on the geographic distribution of seismicity. The first interpretations of seismologic data in terms of plate tectonics revealed the agreement of seismicity with the expected conditions due to the relative motion at plate boundaries (Isacks *et al.*, 1968; McKenzie, 1972).

The number of types of plate boundaries can be reduced basically to three, namely, divergent or *rift zones*, convergent or *subduction zones*, and transcurrent horizontal motion or *transform faults* (Fig. 2.9).

At divergent boundaries and in rift zones, plates separate themselves from one another, and new oceanic lithosphere is created between them, for example, in ocean ridges. For plate boundaries of this type, earthquakes are of shallow depth and of moderate magnitude ($M < 7$), forming a narrow band along boundaries. At convergent boundaries plates collide and one of them is introduced under the other in the mantle in a process called *subduction*. Subduction zones are located in zones of deep earthquakes under orogenic belts or island arcs. The subducted lithosphere is always of an oceanic nature and its dip varies from case

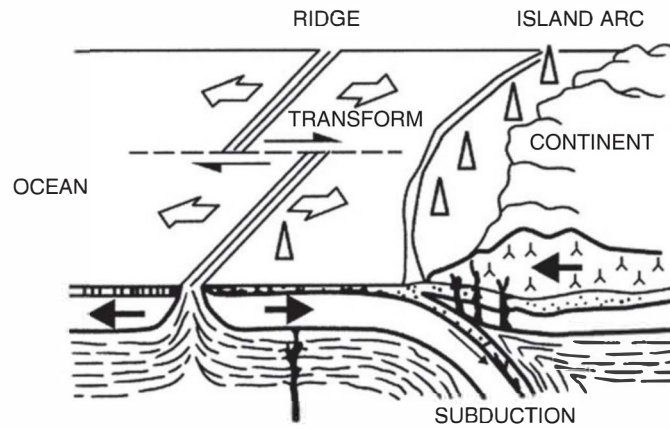


Fig. 2.9

Processes occurring at plate boundaries, namely divergence or extension (at a ridge), convergence or collision (at a subduction zone), and transcurrent motion (at a transform fault).

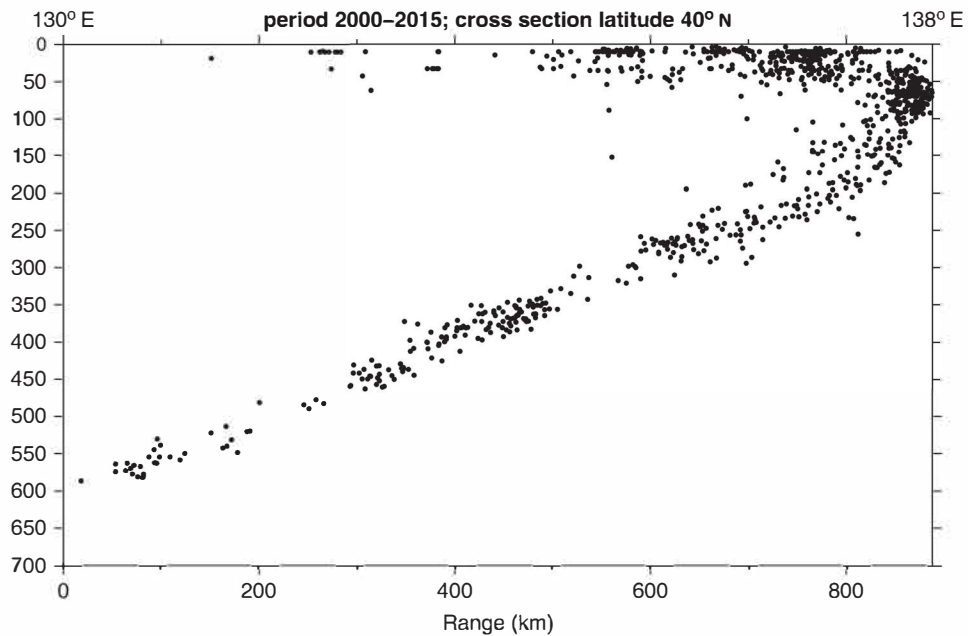


Fig. 2.10

The distribution of earthquakes with depth in the subduction zone of Kurile Islands on a cross-section perpendicular to the coasts (for the period 2000–2015, $M > 5$) (data from NEIC, U.S. Geological Survey).

to case. In subduction zones, subducted plates maintain their rigidity and earthquakes take place from the surface to a maximum depth of 700 km. The seismic zone, called the *Wadati–Benioff zone* (named after Benioff and Kiyoo Wadati), is, generally, limited to the upper part of the subducted plate (Fig. 2.10). Beyond 700 km depth, subducted plates are aseismic and become assimilated into the material of the mantle. If, at a converging

boundary, both plates are continental, processes are more complex, with seismicity extended over a wide area, such as for the boundary between Eurasia and Africa (Fig. 21.1). At convergent boundaries earthquakes reach large magnitudes ($M > 7$). In depth, the upper part of the subducted plate is stretched while the lower part is compressed against the mantle. Extension and subduction zones are connected by boundaries of horizontal transcurrent motion or transform faults, where motion is transformed into extension or compression at their ends (Wilson, 1965). Along transform faults, earthquakes are shallow and may reach large magnitudes ($M > 7$). Two zones of convergence may also be linked by a transform fault. The relation between earthquakes and tectonics will be seen in Chapter 21.

2.8 Earthquake risk, prediction, and prevention

The practical applications of seismology are directed to mitigation of the disastrous effects of earthquakes affecting structures and human lives. As we commented above, this should be the most important concern of earthquake seismologists. In a general way, the risk associated with the occurrence of earthquakes, usually called *seismic risk*, can be separated into two factors: *seismic hazard* and *vulnerability*. A third factor is the economic cost, but this usually falls outside seismologists' interest. The first, seismic hazard, refers to the occurrence of earthquakes affecting a particular area. Hazard can be approached from a *deterministic* or a *probabilistic* point of view, as we will see in detail in Chapter 21. The second, vulnerability, represents the amount of damage that a certain structure may suffer due to the action of an earthquake. Both factors, hazard and vulnerability, are included in the general term, seismic risk. In this form, we separate the purely seismological aspect of seismic hazard from engineering considerations regarding the vulnerability of different structures.

The seismic hazard of a site is given in terms of the ground motion (displacement, velocity, and acceleration) produced by nearby earthquakes. Commonly, only accelerations are considered, and, in many instances, ground motion is represented by the value of the seismic intensity. As we shall see in Chapter 21, the seismic hazard of a site depends on many factors, namely, the distribution of nearby seismic sources, their magnitudes, depth, and focal mechanisms, the nature of the propagating medium, especially wave attenuation and geometric spreading, and the characteristics of the shallow layers under the site that increase or decrease the amplitudes of ground motion.

Adequate assessment of seismic hazard is fundamental to preventing the damage caused by earthquakes, since buildings and other structures can be designed to withstand without collapsing the maximum ground acceleration expected for a given site. Study of the structural responses of various types of buildings is fundamental for safe building practice in seismic regions and is the main concern of *seismic engineering*. All countries with an appreciable seismic risk have *antiseismic building codes* that are sanctioned by law. Safety norms are stricter for structures whose collapse can produce large catastrophes, such as, for example, nuclear plants, dams, hospitals, and other public buildings.

The problem of predicting earthquakes has always been on the horizon of seismology (Simpson and Richards, 1981; Rikitake, 1982; Lomnitz, 1994, Kanamori, 2003, Hough, 2010). For many years, this was considered to be a task beyond the possibilities of true science (Macelwane, 1946). From about 1970 until 1980, attitudes were more optimistic and a solution to the problem was considered to be almost at hand. More recently, a more critical view has returned, recognizing the difficulty of the problem and even hinting at its impossibility (Geller, 1997). However, since forecasting the occurrence of earthquakes with sufficient advance warning would be a very efficient way of diminishing the number of casualties and preventing, in part, some of the damage, it remains an indispensable task for seismologists.

The problem of prediction is based on the identification of precursor phenomena that indicate the impending occurrence of an earthquake. Precursors are observables that are related to changes in physical conditions in the focal region. Strictly speaking, some phenomena, such as the abnormal behavior of animals, cannot be considered true seismologic precursors. Although there have been reports since antiquity that certain animals had sensed sometime before that an earthquake was going to occur, it is difficult to verify and quantify such phenomena, let alone establish their relation to future impending earthquakes. We will examine some of the proposed precursors in Chapter 21. At present, none of the proposed precursors can be considered to be a clear indication of the occurrence of a future earthquake and in consequence we cannot yet speak of a proven method (Wyss and Booth, 1997).

If the prediction of earthquakes remains a distant and problematic question, prevention of damage is perfectly achievable with the existing state of our knowledge. If we cannot predict a future large earthquake, we can mitigate or even prevent in greater part the damage that it will cause, especially in terms of human casualties, with adequate building practice. Recent earthquakes in countries with strict antiseismic design practice and in those without such practice have produced very different numbers of casualties. For example, the Loma Prieta, California, earthquake of 1989 ($M = 7$) produced only 62 fatalities whereas the smaller Latur, India, earthquake of 1993 ($M = 6.3$) produced an estimated 11,000 fatalities. Good assessment of the seismic risk and adequate building practice can avoid the total collapse of buildings and decrease drastically the number of human casualties. We must not forget that it is buildings that produce casualties, not earthquakes *per se*. Appropriate seismic risk assessment and antiseismic building norms will prevent large human disasters such as those we have experienced in the recent earthquakes of Haiti and Nepal. A new tool in prevention are the earthquake early warning systems (EEWS). The EEWS allows warnings to be issued to a site with a short lead-time about the impending arrival of the largest strong ground motion from an earthquake after the first wave arrivals have been detected nearer to the source by adequate sensors (Chapter 21).

2.9 Summary

Large earthquakes are first of all natural disasters that very seriously affect human lives. This aspect should not be forgotten. A first approach to their study is a descriptive one. This

has led to a first classification of their size in terms of their intensity based on the damage caused. Large earthquakes cause important damage and loss of human lives, while very small ones are only felt by sensitive instruments. Earthquakes are caused by rupture on faults which for large ones may extend to a thousand kilometers with displacements of above ten meters. Seismicity is the study of their location, size, and time occurrence. Large earthquakes are generally preceded by foreshocks and followed by a long series of aftershocks.

Study of the propagation of seismic waves gives information about the structure of the Earth. From this information we have found that from the surface to its center the Earth is formed by the crust of thickness about 30 km, the mantle, divided into upper (down to 700 km) and lower to the boundary of the core at 2900 km, and the core divided into outer and inner core.

Study of the occurrence of earthquakes has contributed to the establishment of the modern theory of plate tectonics which explains the dynamics of the Earth's surface. The Earth's surface is divided into plates which move relative to each other causing at their boundary most of the earthquakes. At the boundaries, plates separate, collide, or slide horizontally with respect to each other. Finally, a most important subject is the assessment of seismic hazard and risk which is fundamental to preventing damage caused by earthquakes. Unlike prevention, prediction, however, remains an unsolved problem.

As we have seen in [Chapter 2](#), earthquakes produce motions in the Earth due to the propagation of seismic waves that result in destruction and damage to buildings and other structures with large loss of human life. These effects are the first type of observations about earthquakes, and are often called *macroseismic data* because they are produced by large motions. This is the only information we have for historical earthquakes. From the end of the nineteenth century, ground motion itself began to be measured and recorded by instruments called *seismographs* at different locations on the Earth's surface; which constitute the second type of observations. These observations represent the quantitative measurement of the ground displacement $u(t)$, velocity $v(t)$, and acceleration $a(t)$ as a function of time, which are present in equations of the application of the principles of mechanics that we will consider in the following chapters. For this reason, we first need to know basic ideas about seismographs and of the processing and analysis of their recordings or *seismograms*. The discipline that studies the principles and characteristics of seismographs is called *seismometry*. We will present firstly the basic theory of the seismograph and some examples of analog instruments, to focus later on digital data from *broad-band seismographs*, which are the instruments most in use at present. An important part of this chapter is dedicated to conversion of the different data formats and the use of *SAC* software (Seismic Analysis Code) to process digital data (Tapley and Tull, 1992). The SAC macros used in this chapter to process seismograms may be found in the online resources, together with the digital seismograms, in order for students to be able to reproduce the examples shown in this chapter.

3.1 The historical evolution of seismographs

The oldest instrument used to detect the occurrence of an earthquake was probably constructed in China during the second century AD and is attributed to the philosopher Chian-hen. This instrument consisted of a bronze figure of a dragon with eight heads whose mouths each contained a ball. Inside the figure there was some kind of pendular device that pushed the balls and made them fall when it was shaken by an earthquake. The figure was oriented in the geographic directions so that, upon arrival of the seismic waves, the corresponding ball would fall and show the occurrence and orientation of the shock (Bolt, 1978). In Europe, the first instrument was a mercury seismoscope designed by Jean de la Haute-Feuille in 1703, and consisting of a vessel with mercury connected by eight channels to eight cavities. Earthquakes will make the mercury flow into one or several of the cavities, indicating their

orientation and size (quantities of mercury spilled). It is not certain that the instrument was actually built, although we have a description of it, but similar instruments were built in 1784 by Atansio Cavalli and in 1818 by Niccolò Cacciatores (Ferrari, 1992). Vertical and horizontal pendulums began to be used in around 1750, in Italy by Nicolas Cirillo, Domenico Salsamo, and Duca de la Torre (Dewey and Byerly, 1969). Some of these instruments had an alarm clock to indicate the occurrence of an earthquake or a stylus attached to the mass that left a mark on sand or on a smoked plate of glass, in which case they were then called *seismoscopes*. Luigi Palmieri, in 1859, was probably the first to use the term *seismograph* for an instrument that recorded arrival time and duration of motion (Ferrari, 1992).

The first true seismographs, with continuous recording on a rotating drum using smoked paper on which time marks were also recorded, were designed at the end of the nineteenth century, mostly in Italy, by Fillipo Cecchi, Giovanni Agamennone, Cancani, Giulio Grablovitz, and Giuseppe Vicentini, and are known by their names. These instruments were horizontal and vertical pendulums with large masses and magnifications under 100. In 1890, Milne introduced a new type of seismograph with an inclined pendulum, which, with a very limited length and mass, had a relatively large natural period. Later, in 1915, in collaboration with John J. Shaw, Milne produced the Milne–Shaw seismograph with a mass of 0.5 kg, a period of 8 s, and a magnification of about 200. Similar instruments were designed in Japan by Omori and in Russia by Igor V. Nikiforov. Toward 1900, Wiechert developed, in Germany, a horizontal seismograph consisting of an inverted pendulum that recorded two components with a single mass, and a vertical seismograph with a mass suspended from a spring. Both types of instrument had masses in the range 80–1000 kg, periods of about 12 s, viscous damping, and magnifications in the range 100–1000. Another mechanical seismograph that found extended use was also developed in Germany by Carl Mainka, with masses in the range 200–500 kg and a magnification of about 300. These were purely mechanical instruments with viscous damping, in which the amplification was produced by a system of levers, and recorded on a drum with smoked paper, together with time marks from a clock. Wiechert and Mainka seismographs were very popular and many are still in operation (Dewey and Byerly, 1969).

Significant progress in seismologic instrumentation was made due to the *electromagnetic seismograph*, developed in Russia in around 1910 by Galitzin. This instrument incorporates a coil attached to the mass of the pendulum that moves in the field of a magnet. The electric current generated in the coil by its motion is passed to a galvanometer whose deflection is recorded on photographic paper by a light beam reflected from a small mirror. Damping of the system is provided by the force opposed to the motion of the coil in the field of the magnet. A further development of this instrument in 1920 by Johann Wilip in Estonia resulted in the Galitzin–Wilip seismograph, with 12 s period and a magnification of about 1500. These instruments were the first of a series of electromagnetic seismographs based on a combination of a seismometer and a galvanometer. Another instrument of small dimensions, the torsion horizontal seismograph, was developed by Harry O. Wood and John A. Anderson in 1922 (*Wood–Anderson torsion seismograph*). It is formed by a small mass attached to a metallic fiber that is made to oscillate by torsion and which records on photographic paper via a light beam that is reflected from a mirror attached to the mass. This instrument, with a period of 0.8 s and an amplification of 2800, was used in the first definition of local magnitude by Richter in California.



Fig. 3.1 Photograph of a STS2.5 broad-band seismograph (courtesy of R. Freudenmann, Strekeisen, A. G.).

Toward 1930, Benioff developed the *variable-reluctance seismograph*, the action of which is based on changes in reluctance due to variations in the distance between two magnets, one fixed to the frame and the other to the mass of a pendulum. Owing to their relative motion, a current is generated in a coil around the magnet of the mass that is passed to a galvanometer and is recorded on photographic paper. This instrument, known as the Benioff seismograph, has a period of 1 s and a magnification in the range 1000–100,000. After around 1945, several models of electromagnetic seismographs with short, intermediate, and long periods, were developed by William Sprengnether with the collaboration of Macelwane at Saint Louis. For long periods, Frank Press and Maurice Ewing developed in 1953 a seismometer–galvanometer system with seismometer periods in the range 15–30 s and galvanometer periods of 100 s, which operated with magnifications in the range 750–6000 (*Press–Ewing seismograph*). The combination of short-period and long-period seismographs avoided the microseismic noise present for intermediate periods of around 6 s and was used at many seismographic stations (Melton, 1981). This combination was used in the stations of the WWSSN (World Wide Standard Stations Network), installed by the U.S. Government around the world, which operated between 1960 and 1990. Similar systems of electromagnetic seismographs were developed in the USSR by Dmitrii P. Kirnos, such as those with a seismometer period

of 12.5 s and a galvanometer period of 1.2 s, resulting in a nearly flat response in the range 0.2–10 s, and those of seismometer periods 0.6–1 s and galvanometer periods 0.2–0.4 s, resulting in a peak response in the range 0.2–1 s (Sawarensky and Kirnos, 1960).

Since 1970, three developments have influenced seismologic instrumentation. The first was the *electronic amplification* of the electric signal from the seismometer by means of electronic amplifiers, which increased the possibilities of the seismometer–galvanometer system. The second was the introduction of *digital recordings* that could be analyzed directly by computers. Digital recording is carried out by means of an analog–digital converter that transforms the electric output of the seismometer. The first instrument of this type (*Caltech digital seismograph*) was developed in 1962. The combination of electronic amplification and digital recording is the basis of all modern seismographs. The third was the development of *broad-band seismographs*, that is, systems with a response curve that is practically flat for a wide range of periods, for example 0.1–1000 s. For this purpose a new seismometer that uses a leaf-spring and a feed-back circuit to extend the response to very low frequencies was developed by Erhard Wieland and G. Streckeisen in 1982 (Wieland and Streckeisen, 1982). With instruments of this type, the duality of short- and long-period instruments has been overcome. Broad-band seismographs record digitally and, with adequate filters, can simulate a response in any frequency range. These instruments are widely used in modern seismographic stations (Fig. 3.1).

3.2 The theory of the seismometer

The physical principles of most types of seismometers are based on the forced motion of a pendulum, be it vertical or horizontal. When the ground moves due to the arrival of seismic waves, it produces a displacement of the frame of the pendulum with respect to the mass due to its inertia. This motion, conveniently amplified, is recorded as a function of time in seismograms. From this relative displacement, we can deduce the ground motion. Basically, a seismogram is a time signal $s(t)$ recorded by a seismograph at a particular location at an epicentral distance Δ from the source, corresponding to one component of the motion of the ground from an incoming seismic wave, and can be expressed as the convolution

$$s(t) = u(t) * I(t), \quad (3.1)$$

where $u(t)$ is the ground motion displacement at an observation point (station), and $I(t)$ is the *instrumental response* (the symbol $*$ denotes the convolution operator, A4.3). In this equation the $I(t)$ is known for each instrument type. We will start explaining the basic ideas about the most common types of seismographs and later the processing of seismograms to obtain the ground displacement $u(t)$.

In order to understand the basic principle of the theory of a seismometer, we will consider an ideal system, consisting of a vertical pendulum with a mass m suspended by a spring of elastic coefficient K and a dashpot with viscous damping c (Fig. 3.2). When the frame of the

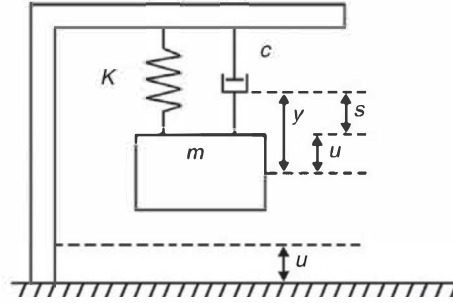


Fig. 3.2

An ideal vertical seismometer formed by a mass hanging from a spring and a dashpot.

pendulum is displaced by $u(t)$ (the ground vertical displacement), the mass moves $y(t)$, and the relative displacement of the mass with respect to the frame is $s(t)$.

With respect to a reference system at rest, since the spring and the dashpot are affected only by the relative motion of the frame and the mass $s(t)$, the equation of motion (Newton's second law) for the mass is given by

$$m\ddot{y} = -Ks - c\dot{s}. \quad (3.3)$$

If we displace the mass with the frame at rest, $y(t) = s(t)$ and the equation of motion becomes, after division by m ,

$$\ddot{s} + \frac{c}{m}\dot{s} + \frac{K}{m}s = 0. \quad (3.4)$$

For a system without damping ($c = 0$), the solution of (3.4) is given by harmonic motion,

$$s(t) = A \sin(\omega_0 t). \quad (3.5)$$

The natural frequency of the undamped system is given by

$$\omega_0 = \left(\frac{K}{m}\right)^{1/2}. \quad (3.6)$$

On substituting in ω_0 , equation (3.4) can be written as

$$\ddot{s} + 2\beta\omega_0\dot{s} + \omega_0^2 s = 0, \quad (3.7)$$

where β is the damping factor,

$$\beta = \frac{c}{2(Km)^{1/2}}. \quad (3.8)$$

If $\beta = 1$, the system is critically damped and $c = 2(Km)^{1/2}$. For a damped system the solution of (3.7) is

$$s = A e^{-\beta\omega_0 t} \sin[\omega_0(1 - \beta^2)^{1/2}t + \varepsilon], \quad (3.9)$$

where A (the amplitude) and ε (the phase) are constants. Equation (3.9) represents *damped harmonic motion* of frequency $\omega_0(1 - \beta^2)^{1/2}$.

If the frame is displaced by $u(t)$, then, on making the substitution $y(t) = s(t) + u(t)$ in (3.4), we obtain

$$\ddot{s} + 2\beta\omega_0\dot{s} + \omega_0^2 s = -\ddot{u}, \quad (3.10)$$

where $\ddot{u}(t)$ is the ground vertical acceleration. The same equation applies to the horizontal components. This equation is called the *seismometer equation*, since it relates the relative motion $s(t)$ of the mass and the frame that can be recorded and measured to the ground motion $u(t)$. If the motion is very fast, the acceleration term is predominant and we have $\ddot{s} = -\ddot{u}$. Then, $s = -u$; that is, the relative displacement of the mass corresponds to the ground displacement. If the motion is very slow ($\ddot{s}$ and $\dot{s}$ are small), then $s = -\ddot{u}/\omega_0^2$; the relative displacement of the mass corresponds to the acceleration of the ground.

If we apply to the frame a harmonic motion of frequency ω , $u(t) = U \sin(\omega t)$, equation (3.10) becomes

$$\ddot{s} + 2\beta\omega_0\dot{s} + \omega_0^2 s = U\omega^2 \sin(\omega t). \quad (3.11)$$

The solution of this equation, without considering the transient part, is

$$s = S \sin(\omega t - \varepsilon), \quad (3.12)$$

where S and ε , the amplitude and phase of the relative motion between the mass and the frame, are

$$S = \frac{\omega^2 U}{[(\omega_0^2 - \omega^2)^2 + (2\beta\omega\omega_0)^2]^{1/2}}, \quad (3.13)$$

$$\varepsilon = \tan^{-1} \left(\frac{2\beta\omega\omega_0}{\omega_0^2 - \omega^2} \right). \quad (3.14)$$

Thus, the amplitude S of the relative motion of the mass depends on the amplitude of the ground motion U , its frequency ω , the natural frequency of the pendulum ω_0 , and its damping β . The motion $s(t)$ measured by the seismometer is shifted by a phase ε with respect to the ground motion $u(t)$. The ratio S/U depends on the frequency and damping (Fig. 3.3). When $\omega = \omega_0$, the system is in resonance and, if it is undamped ($\beta = 0$), the ratio becomes infinite and the phase is $\varepsilon = \pi/2$. For critical damping ($\beta = 1$), when $\omega = \omega_0$, the ratio $S/U = \frac{1}{2}$. For large values of ω , $S/U = 1$ and $\varepsilon = \pi$. Usually, a seismometer has

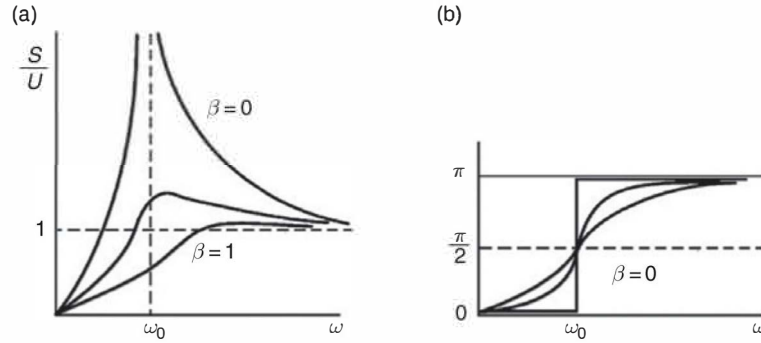


Fig. 3.3 Frequency-response curves of a mechanical seismometer: (a) the amplitude response and (b) the phase response.

a damping near its critical value. Equations (3.13) and (3.14) represent the amplitude and phase responses of a seismometer.

Another way to study this problem is to consider the response of the system to an impulsive acceleration, that is, $\ddot{u}(t) = \delta(t)$, the *Dirac delta function* (introduced by Paul Dirac) (see Section 4.7). By substitution into (3.9), we obtain

$$\ddot{s} + 2\beta\omega_0\dot{s} + \omega_0^2s = -\delta(t). \quad (3.15)$$

If we take the *Fourier transform* (developed by Joseph Fourier) (FT, see Appendix 4), then, since the transform of $\delta(t)$ is 1 and that of $s(t)$ is $S(\omega)$,

$$-\omega^2S(\omega) + 2i\omega\omega_0\beta S(\omega) + \omega_0^2S(\omega) = -1. \quad (3.16)$$

We obtain for $S(\omega)$,

$$S(\omega) = \frac{-1}{-\omega^2 + 2i\beta\omega\omega_0 + \omega_0^2}. \quad (3.17)$$

Using $S(\omega) = |S(\omega)|\exp[i\varepsilon(\omega)]$, we obtain for the amplitude and phase responses two expressions similar to (3.13) and (3.14) (with $U = 1$):

$$|S(\omega)| = \frac{1}{[(\omega_0^2 - \omega^2)^2 + (2\beta\omega\omega_0)^2]^{1/2}} \quad (3.18)$$

$$\varepsilon(\omega) = \tan^{-1}\left(\frac{2\beta\omega\omega_0}{\omega_0^2 - \omega^2}\right). \quad (3.19)$$

The response in time $s(t)$ is found by taking the inverse transform of $S(\omega)$. The response of the seismometer to an impulsive acceleration represents its behavior for all frequencies. The response for an acceleration of arbitrary form $\ddot{u}(t)$ can be obtained by its convolution with the response to the impulsive acceleration or, in the frequency domain, by taking the product of their transforms.

3.3 Recording systems, magnification, and dynamic range

The motion of the mass of the seismometer is recorded in analog or digital form after its amplification. The complete system of the seismometer, amplifier, and recorder is known as a *seismograph*. The first seismographs were totally mechanical systems in which amplification was carried out by means of systems of levers, and recording was by a stylus onto smoked paper fixed onto a revolving drum. This system is today obsolete, but it is useful to study in order to understand the behavior of a seismograph.

The ratio S/U in (3.13) is called the *dynamic magnification* $V_d(\omega)$ of the seismometer for a harmonic input signal of frequency ω :

$$V_d(\omega) = \frac{S}{U} = \frac{\omega^2}{[(\omega^2 - \omega_0^2)^2 + 4\omega^2\omega_0^2\beta^2]^{1/2}}. \quad (3.20)$$

The frequency *transfer function* $T(\omega)$ is defined as the FT of the output signal divided by the FT of the ground motion or input signal (Scherbaum, 1996, pp. 63–65),

$$|T(\omega)| = \frac{S}{U} = \frac{\omega^2}{[(\omega_0^2 - \omega^2)^2 + (2\beta\omega\omega_0)^2]^{1/2}} \quad (3.21)$$

and the phase is given by eq. (3.19).

If instead of the FT we apply the *Laplace transform* (developed by Pierre Simon Laplace) (LT, see Appendix 4) to the seismometer equation (3.10), we obtain (since in LT it is customary to use the variable $s = c + i\omega$, here we call the relative motion Z)

$$\omega^2 Z(s) + 2\beta\omega_0 s Z(s) + \omega_0^2 Z(s) = -\omega^2 U(s) \quad (3.22)$$

and the transfer function is now given by

$$T(s) = \frac{Z(s)}{U(s)} = \frac{-s^2}{s^2 + 2\beta\omega_0 s + \omega_0^2}. \quad (3.23)$$

In terms of the LT, the transfer function, given by equation (3.23), is characterized by its *zeros*, the roots of the numerator, and by its *poles*, the roots of the denominator. Thus equation (3.23) has two zeros (numerator equal to 0) and two complex conjugate poles, the roots of the polynomial in the denominator.

In mechanical seismographs, by using a system of levers we can amplify the amplitude of the relative motion of the mass, so that the recording amplitude is, $S' = V_s S$, where the factor V_s is called the *static magnification*. The total magnification of the system is given by the product of both factors, $V(\omega) = V_s V_d(\omega)$. The total response of the seismograph is given by the *magnification curve* in the frequency domain $V(\omega)$ which, according to (3.13) and (3.14), is formed by the amplitude and phase responses. The maximum value of the amplitude magnification V_{max} and its corresponding period are used as characteristics of

the response of each seismograph. The ground motion $u(t)$ is obtained from the recorded signal $s'(t)$, dividing its transform by the response of the instrument:

$$U(\omega) = \frac{S'(\omega)}{V(\omega)}. \quad (3.24)$$

To obtain the time function $u(t)$ of the ground motion, we take the inverse transform of $U(\omega)$ (A4.11). An approximation to the amplitude of the ground motion for a given period can be obtained by dividing the recorded amplitude by the corresponding magnification for that period.

Mechanical seismographs are very limited in amplification by friction between their parts, and the dimensions of the pendulum and recording systems. To increase magnification, their mass is increased in order to overcome friction, reaching several tons in some cases. The Wiechert seismograph of 1000 kg had a maximum magnification of nearly 1000 and the Mainka of 350 kg had magnification of about 400. A recording system avoiding friction between the pen and paper uses a light beam and photographic paper, such as in the Wood–Anderson device that, with a small mass of a few grams, attained a magnification of 2800.

In the problem of the amplification of signals, an important concept is the *dynamic range*, that is, the range between the maximum and minimum amplitudes possible for a particular system. The dynamic range of a system is given by the logarithm of the ratio A/A_0 , where A is the maximum amplitude recorded and A_0 is the minimum amplitude or that taken as the zero level. The units used are *decibels* (dB) such that, for a given ratio, its dynamic range is $20 \log(A/A_0)$ dB. For example, if A/A_0 equals 1000, the dynamic range is 60 dB. For a seismograph, A and A_0 are the maximum and minimum recorded amplitudes. In an analog record, A_0 is related to the minimum detectable amplitude and the noise generated by the system itself, and A is related to the dimensions of the record or the *saturation level* of the recording system. The saturation level is the maximum amplitude possible with a particular recording system. For example, for a photographic analog seismogram, taking into account the thickness of the trace and the size of the record, the minimum appreciable amplitude is about 0.5 mm and the maximum is 20 cm; thus, $A/A_0 = 1000$ and the dynamic range is 60 dB.

The dynamic range is in itself independent of the magnification. If we increase the magnification for the same dynamic range, the system is saturated for smaller amplitudes of ground motion and we lose information about large amplitudes. For example, if the maximum amplitude of a graphical record is 10 cm and the minimum is 1 mm, its dynamic range is 40 dB. If the maximum magnification is 10,000, the minimum ground motion recorded is 0.1 μm and the maximum is 10 μm . If we increase magnification to 100,000 we can detect ground motions of 0.01 μm amplitude, but the system saturates for amplitudes larger than 1 μm .

3.4 Electromagnetic seismographs

The foundation of the electromagnetic seismograph consists in adding a coil to the mass of the pendulum that moves in the magnetic field of a magnet. The same effect is produced if

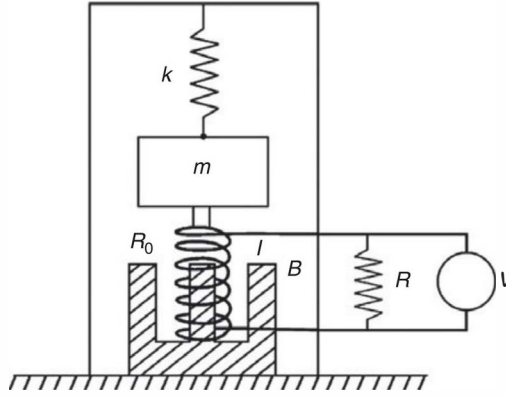


Fig. 3.4 A vertical electromagnetic seismometer.

the moving part is a magnet inside a fixed coil. In both cases, the relative motion of the coil in the magnetic field generates an electric current in the coil that is proportional to the velocity of the relative movement of the coil and the magnet. This electric current is passed to a galvanometer whose deflection is recorded graphically.

Let us consider the same ideal case of a vertical pendulum as that in Section 3.3, with a moving coil and a fixed magnet (Fig. 3.4). The force F that acts on the mass of the pendulum due to the motion of the coil in the magnetic field B of the magnet, according to the Biot–Savart law (after Jean Baptiste Biot and Félix Savart), is

$$F = IBl, \quad (3.25)$$

where I is the current in the coil, $l = 2\pi rN$ is the coil length (r is the radius and N is the number of turns), and B is the amplitude of the magnetic induction. If s is the relative displacement of the coil and the magnet, the work done by the force is Fs and the power is $F\dot{s}$. In the electric circuit (Fig. 3.4), power is dissipated by its total resistance $R_0 + R$ (R_0 is the coil's resistance and R is an additional resistance in parallel). The power in the circuit is given by IV , where V is the difference in electric potential, thus

$$VI = I\dot{s}. \quad (3.26)$$

If we make $G = IB$, a constant that depends on the specifications of the instrument, we have that $V = G\dot{s}$ and $F = GI$. Using Ohm's law (after Georg Simon Ohm), $I = V/R$, we obtain

$$I = G \frac{\dot{s}}{R_0 + R}, \quad (3.27)$$

$$F = G^2 \frac{\dot{s}}{R_0 + R}. \quad (3.28)$$

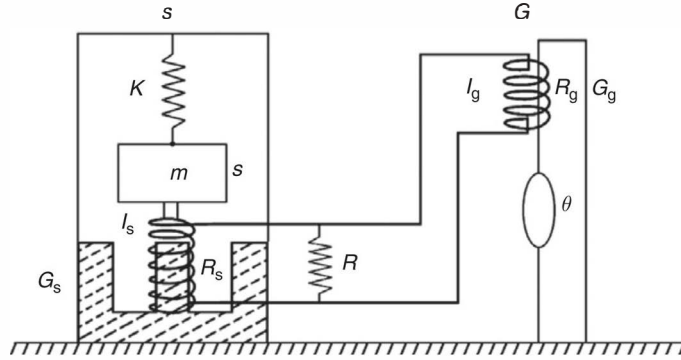


Fig. 3.5

An electromagnetic seismograph formed by a vertical seismometer and a galvanometer.

The force that acts on the mass due to the motion of the coil in the magnetic field depends on its velocity and acts in the opposite sense, that is, it is a damping force. The resulting equation of motion has the same form as (3.8), but now the damping coefficient is

$$2\beta = \frac{G^2}{(Km)^{1/2}(R_0 + R)}. \quad (3.29)$$

The damping of the system is electromagnetic and its critical value ($\beta = 1$) corresponds to

$$\frac{G^2}{R_0 + R} = 2m\omega_0. \quad (3.30)$$

The damping can be adjusted changing the value of the resistance in parallel R , for each fixed value of the resistance of the coil R_0 .

The recording of the electromagnetic seismograph results from connecting the current generated in the coil of the seismometer to a galvanometer whose angular deflection is recorded on a drum with photographic paper by using a beam of light that is reflected from a mirror attached to the galvanometer's moving part (Fig. 3.5). Since the seismometer and the galvanometer are connected, the current in the circuit generated by the motion of the mass is affected by the motion of the galvanometer. The motion of the mass of the seismometer is also affected by a force that depends on this current. The motions of the galvanometer and of the mass of the seismometer are coupled through the circuit, so that their equations must be solved together. The equations of motion of the seismometer and galvanometer are

$$\ddot{s} + 2\beta_s \omega_s \dot{s} + \omega_s^2 s = -\ddot{u} - G_s I_s / m, \quad (3.31)$$

$$\ddot{\theta} + 2\beta_g \omega_g \dot{\theta} + \omega_g^2 \theta = G_g I_g / h, \quad (3.32)$$

where s is the relative vertical displacement of the seismometer and θ is the angular deflection of the galvanometer. The subindexes s and g refer to the seismometer and

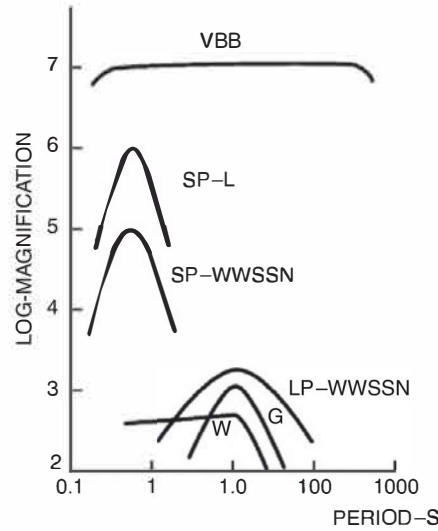


Fig. 3.6

Amplitude-response curves of Weichert (W), Galitzin (G), SP-WWSSN, LP-WWSSN, short period for local seismicity (SP-L), and broad-band (VBB) seismographs.

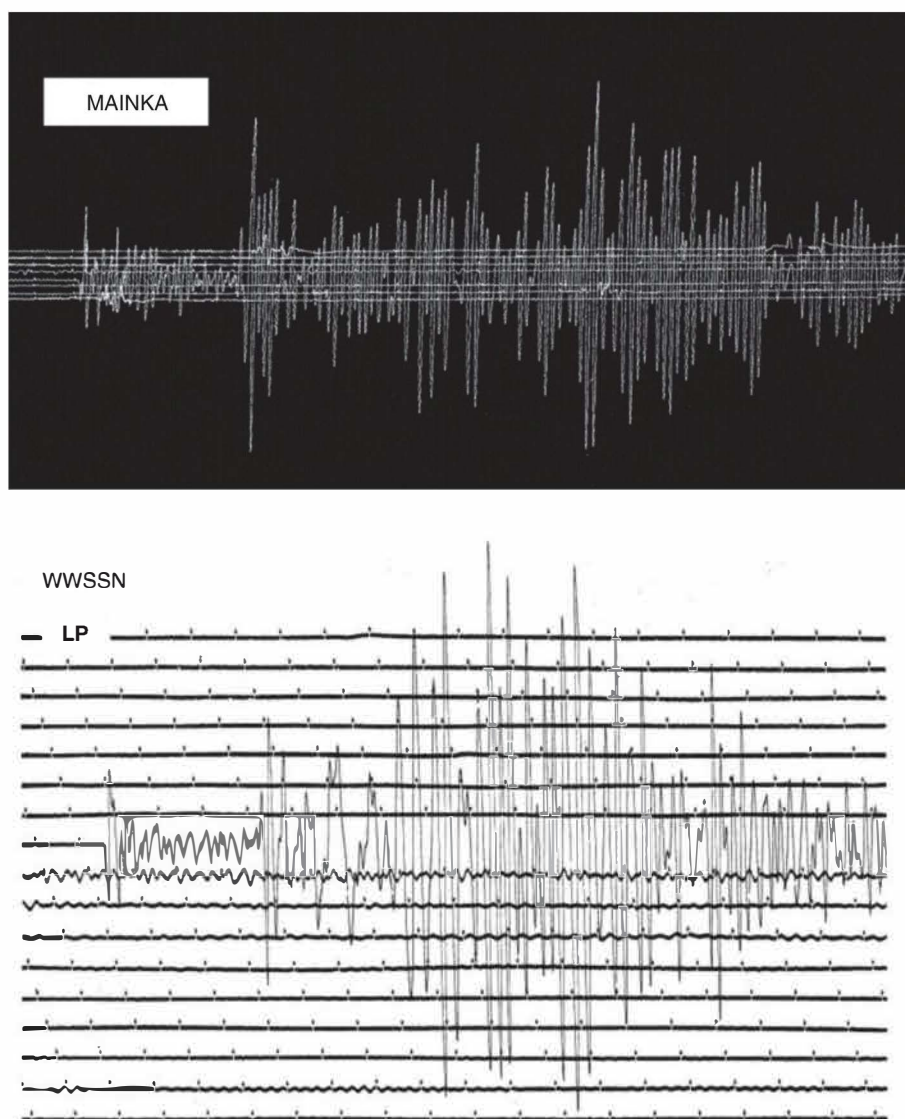
galvanometer, and h is the inertial moment of the moving part of the galvanometer. The electric currents that pass through the coils of the seismometer and galvanometer are interrelated. The recorded motion $\theta(t)$ due to a ground motion $u(t)$ is obtained by the solution of a fourth-order differential equation. In equations (3.31) and (3.32), ω_s and ω_g are the natural frequencies of the seismometer and galvanometer whose values are characteristic to each system.

The seismometer-galvanometer system responds to a ground displacement $u(t)$ with a deflection of the galvanometer $\theta(t)$, which is recorded on photographic paper. Since the current generated in the moving coil of the seismometer is proportional to its relative velocity with respect to the magnet fixed to the ground, the instrument records the ground velocity $\dot{u}(t)$. The total magnification of the system is given by

$$V(\omega) = V_{st}V_sV_g, \quad (3.33)$$

where V_{st} is the static magnification described by the mechanical seismograph, and V_s and V_g are the dynamic magnifications of the seismometer and galvanometer, which are obtained from the solutions to equations (3.31) and (3.32). The result is expressed in terms of the response curves for the amplitude and phase relating the recorded motion to the ground displacement.

A later development of electromagnetic seismographs was the inclusion in the circuit of an electronic amplifier and replacement of the galvanometer by a transducer that transforms the amplified current into the motion of a recording pen. A very common type of recorder involves a heated pen that leaves a mark on thermosensitive paper. Modern seismographs include operational amplifiers of various types, with optional filters. In this way, in theory, the magnification of the instrument can be made as high as desired.

**Fig. 3.7**

Top: A seismogram from the Mainka seismograph (mechanical) of the EBR station (Observatorio del Ebro, Spain) for the earthquake of 9 February 1948 in Greece. Bottom: A seismogram of the vertical long period seismograph (WWSSN) (electromagnetic) of the TOL station (Observatorio de Toledo, Spain) for the 17 September 1972 earthquake in Greece.

As we have seen, mechanical seismographs record ground displacement. Their seismograms are still used for early-twentieth-century earthquakes (Fig. 3.7 top). In about 1920 electromagnetic seismographs were introduced and rapidly substituted mechanical ones. Their seismograms, in analog or digital form, represent the velocity of ground motion (Fig. 3.7 bottom).

3.5 Digital seismographs

Most instruments that we have described previously, such as Weichert and Mainka (mechanical) or Galitzin and WWSSN (electromagnetic), are *analogical seismographs* and provide graphical *analogical records* (seismograms), corresponding to the three components of ground motion (vertical, NS, and EW). These instruments were used until 1980 when *digital instruments* began to be installed. Analogical records are continuous, so that before we can use them in digital computers, which began to be used in seismology in around 1960, they must be digitized. The digitization process has a number of limitations, such as the length of sampling interval, which depends on a graphical representation of the seismogram, and errors which are easily introduced during the process. Other problems with analogical seismographs are the low dynamic range and that their response curve is centered at a narrow frequency range. For example, the WWSSN network had two types of instrument: short period (SP) centered at 1 s and long period (LP) at 15 s (Fig. 3.6). Digital instruments change the continuous electric current output of an electromagnetic seismograph into a digitized time series using an *analog-to-digital* (AD) converter. The sampling rate can be as high as needed, depending on the AD converter. At present, most seismographic stations use digital seismographs and provide data in digital form. Another modern improvement in the instrumentation has been substitution of the short period (SP) and long period (LP) of analogical instruments by digital *broad-band* instruments (BB) that have a flat response between a frequency of 20 Hz and a period of 1000 seconds or 1 mHz (Fig. 3.6). Thus, one single instrument substitutes for the previous two. Because of the generalization of this new type of instrument, we will limit our presentation to the analysis of data from digital broad-band seismographs.

Normally, digital broad-band instruments are provided with different channels with different sampling time intervals and frequency bands; the most common currently in use are listed in Table 3.1.

Amplitudes of digital seismograms are given in *counts* or *digital units* (DU). The equivalence between counts and displacements or velocities depends on the instrument amplification (a common equivalence is, approximately, 1 count = 10^{-9} m or 1 nm for displacement output, or 10^{-9} m/s or 1 nm/s for velocity output). The amplitude instrumental magnification of displacement is given by equation (3.20). In digital instruments the static

Table 3.1 Broad-band instrument channels and sampling rates.

Channel	Sampling rate (samples/s)	Time interval (s)
VH	1/10	10
LH	1/1	1
BH	20/1	0.05
HH/SH	50/1–100/1	0.02–0.01

magnification V_s is given by the *instrument constant*, the product of two terms: A_o (*normalization factor*), given in $V/(m/s)$, and the *sensitivity* D_s , given in V/count . As a consequence, the total instrument response $I(s)$ in the most general form, including the AD converter and filters, is given in the form of a product of poles and zeros, as a function of $s = i\omega$, by

$$I(s) = A_o D_s \frac{\prod_{j=1}^{N_z} (s - s_z)}{\prod_{j=1}^{N_p} (s - s_p)}, \quad (3.34)$$

where N_z is the number of zeros and N_p the number of poles, and s_z and s_p are their respective values.

As digital instruments provide channels with different sampling rates (Table 3.1), the data selection depends on the problem that we need to solve. For example, if we only want to pick the arrival times of different phases at regional distances (less than 500–1000 km) from small earthquakes we select the HH channel, but if we want to obtain the group velocity of surface waves at teleseismic distances, the LH channel is enough for this purpose. Another important point is to remember that while mechanical seismographs record ground displacement $u(t)$, electromagnetic instruments (digital instruments are electromagnetic) record ground velocity $v(t)$.

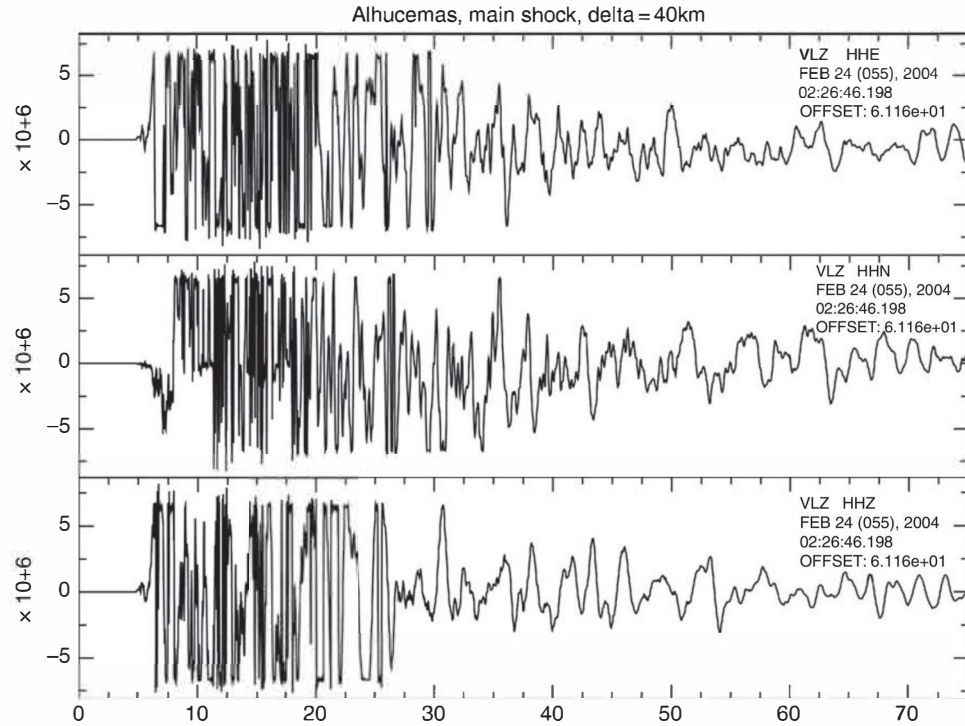


Fig. 3.8

Clipped seismograms at broad-band station VLZ (Velez Gomera, North Africa, Western Mediterranean Network) for the Al Hoceima earthquake of 24 February 2004, $M_w = 6.4$, at 40 km epicentral distance. Figure shows the east–west (HHE), north–south (HHN), and vertical components for the HH channel.

In spite of their large dynamic range, digital broad-band seismographs also *saturate* and *clip* with ground motion of large earthquakes at short distances, making their data unusable. For example, an $M_w = 6.4$ earthquake recorded by a broad-band instrument at 40 km epicentral distance will be saturated (Fig. 3.8). To avoid this problem we can use other instruments to study large earthquakes in the near field, such as *strong motion* instruments (*accelerographs*) or GPS (the main characteristics of these instruments will be explained in Section 3.9)

3.6 Processing digital seismograms

Processing a digital seismic signal, which involves a series of processes starting with data acquisition, is a necessary first step prior to any seismological analysis. Here we will present some of these processes: (1) determination of the radial and transverse component of horizontal motion; (2) instrumental deconvolution or removal of the instrument to obtain ground motion; (3) calculation of spectra; (4) filtering a seismic signal. To process the seismograms we use SAC software. The seismograms and macros used in the examples presented are given in the electronic material. Details about SAC software may be found in the SAC manual and Helffrich *et al.* (2013).

3.6.1 Data acquisition

Digital seismograms from broad-band stations can be obtained from several websites for data centers, for example, IRIS (www.iris.edu/hq), GEOFON (<http://geofon.gfz-potsdam.de/geofon/>), and ORFEUS (www.orfeus-eu.org). Data on these websites correspond to moderate/large earthquakes ($M_w \geq 4.5$ –5.0). Data are given in binary form; one of the most commonly used formats is the SEED (Standard for the Exchange of Earthquake Data) format. From these websites one can also download the software necessary to process seismograms in SEED format. In the IRIS website the program is called *rdseed* (the complete instructions to run this program are in the electronic material (EM_3_1)). As this program is updated periodically, one should be careful to download the latest version.

As an example we will use the record of the Nepal earthquake of 12 May 2015 recorded at the BHZ channel of station UCM (Western Mediterranean Network, Universidad Complutense de Madrid, Orusco de Tajuña, Madrid, Spain). The input for the *rdseed* program is the file WM.UCM..BHZ.D.2015.132. This input file is given in miniSEED format, a seed file that does not have the header file which contains information about the instrument response. In this case, if we want to have the instrument response we have to execute before the following command: `setenv ALT_RESPONSE_FILE dataless_sta`, where `dataless_sta` is a header file that contains the station characteristics, such as station coordinates, instrument type, etc.

The output consists of three files (see electronic material). The first, 2015.132.00.00.16.9950.WM.UCM..BHZ.D.SAC, is the seismogram in SAC format, the other two files are the instrument response given in terms of a Laplace transform: the file RESP.WM.UCM..BHZ

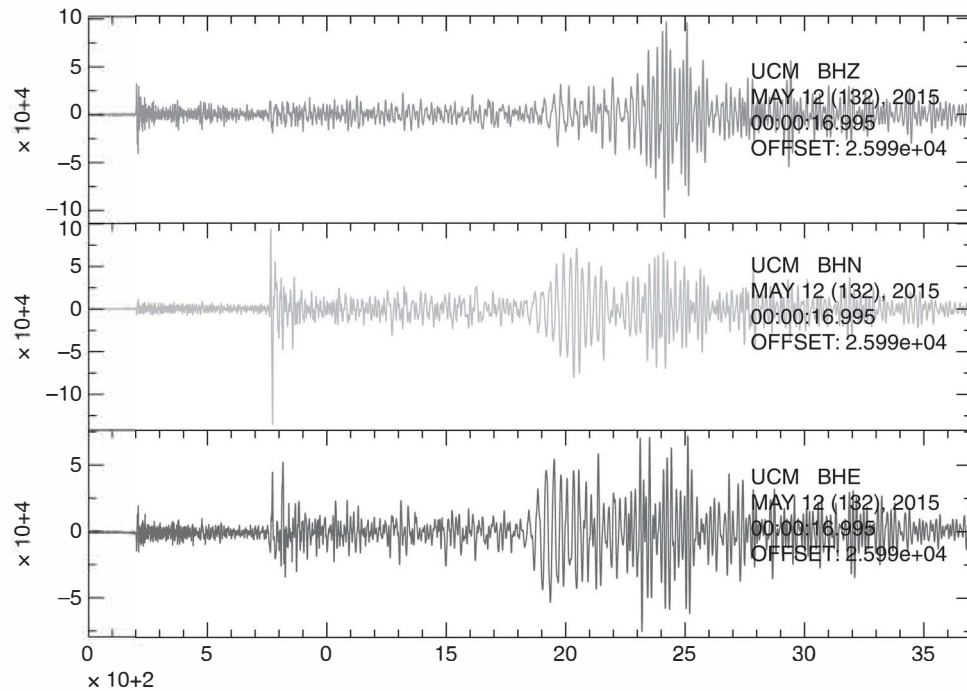


Fig. 3.9

Seismograms of the three components (BHN north–south, BHE east–west, and BHZ vertical) of the UCM station (Western Mediterranean Network, Universidad Complutense de Madrid, Orusco de Tajuña, Madrid, Spain) for the Nepal earthquake of 12 May 2015. (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

is the velocity instrument response and SAC_PZs_WM_UCM_BHZ__2010.173.00.00.00.0000_2599.365.23.59.59.99999 is the displacement instrument response (see [Section 3.6.3](#) for more details). The formats that can be retrieved with the *rdseed* program are: GSE (Group Scientific Experts), CSS (Center for Seismic Studies), SEISAN, etc. We can do the conversion from these formats to the SAC format, for example, for the CSS format, using the *css2sac* command, but we must include a file with the station's characteristics. If we are not interested in the instrument response, we can use the command *mseed2sac* to obtain the seismogram in SAC format, but in this case we do not get the instrument response. Seismograms can be read using other software, such as, for example, MATLAB. Finally, the seismogram data are plotted as shown in [Fig. 3.9](#).

3.6.2 Radial and transverse components

If we want to obtain the *radial* and *transverse* components of horizontal motion to separate P, SV from SH in body waves ([Chapter 6](#)) and Rayleigh from Love in surface waves ([Chapter 12](#)), we have to rotate the horizontal components (NS and EW). Some instruments are not oriented exactly in the NS and EW directions, but the horizontal components are rotated some degrees to the north and east and called BH1 and BH2. Using the SAC

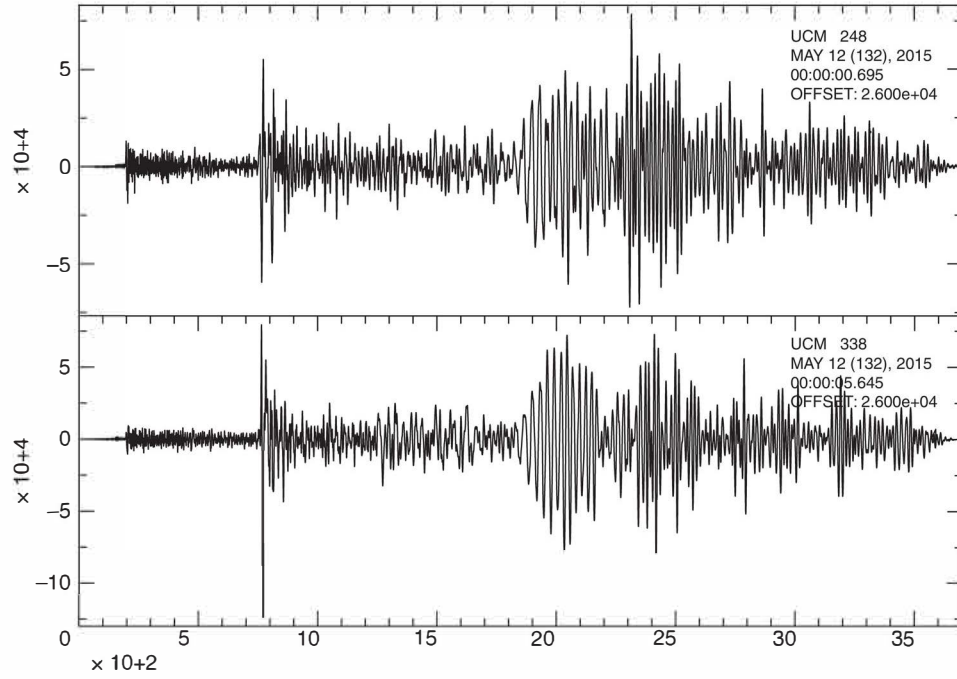


Fig. 3.10 Radial (top) and transverse (bottom) horizontal components of the seismogram of Fig. 3.9.

command *lh* we can retrieve this information. To obtain the radial and transverse components for the record of the Nepal earthquake at the UCM station, we use the *rotate* command to obtain the horizontal radial and transverse components (EM3_2.m). We firstly read the data file in this order: first the NS component and then the EW component; we then remove the mean and the trend (*rmean* and *rtrend* commands) and we use the *rotate* command to obtain the radial and transverse components. Finally, we plot the result in graphic mode or obtain a ps file (*bd sgf* command). The result is shown in Fig. 3.10. To use the *rotate* command we need both files to have the same number of points and the latitude and longitude of the station and event to be included at the header of the SAC file. If we do not have this information we can add it using the *chnhdr* SAC command.

3.6.3 Removing the instrumental response

An important item in seismogram analysis is to obtain the true ground motion from the seismogram. Using equation (3.1) the ground displacement $u(t)$ can be found if we know the instrument response $I(t)$ but, as we have previously mentioned, most seismographs are velocity instruments, that is, they record the ground velocity $v(t) = \dot{u}(t)$. The instrumental response for this kind of instrument is a velocity response $I_v(t)$, so that we modify equation (3.1) into the form

$$s(t) = v(t) * I_v(t). \quad (3.35)$$

If we apply the FT to this equation, we obtain in the frequency domain

$$S(\omega) = V(\omega)I_v(\omega). \quad (3.36)$$

And using the derivative property of the FT in this equation we can write

$$S(\omega) = i\omega U(\omega)I_v(\omega). \quad (3.37)$$

Solving for the FT of the ground displacement we get

$$U(\omega) = \frac{S(\omega)}{i\omega I_v(\omega)}. \quad (3.38)$$

Taking the inverse FT in equation (3.38), we obtain the ground motion displacement in the time domain $u(t)$. In order to obtain the ground displacement from equation (3.37), we remove the effect of the instrument, using the velocity instrument response $I_v(t)$, and afterwards integrate this equation to obtain $u(t)$. Another way to do this is using the LT in terms of poles and zeros (3.34), adding a zero to the instrumental response and thus obtaining directly the ground displacement $u(t)$. This second option is more convenient because it implies fewer calculations. Adding a zero to a velocity instrumental response is equivalent to using the displacement instrumental response,

$$I_d(s) = sI_v(s). \quad (3.39)$$

These two ways to remove the instrument explain the two different files for the instrument response obtained in Section 3.6.1, namely: RESP.WM.UCM..BHZ (velocity instrument response) and SAC_PZs_WM_UCM_BHZ__2010.173.00.00.0000_2599.365.23.59.59.99999 (displacement instrument response). The difference between the two files is the number of zeros: two zeros for velocity and three zeros for displacement, according to equation (3.39). In Fig. 3.11 we have the original seismogram for the Nepal earthquake recorded at station UCM, component BHZ, and the ground displacement record (called ucm_dec.bhz a SAC file). The macro to obtain this figure is given in the electronic material (EM3_3.m). It is important to remember that each component has its own instrument response which may be different. Thus, if we want to obtain the radial and transverse components of the ground motion displacement, previously to rotating the components, we have to remove the instrument using the instrument response for each component and do the rotation after.

3.6.4 Spectral analysis

Spectral analysis of seismic waves is a fundamental tool in seismology. We can obtain the *amplitude* and *phase spectra* of the ground motion $u(t)$ by applying the FT. The macro to perform the FT is given in EM3_4.m. In our example of the Nepal earthquake recorded at the UCM station, to determine the spectra of the ground motion displacement, we use the file ucm_dec.bhz, obtained previously (see Section 3.6.3), where we have removed the instrument. First we select a certain time interval of the record (in our case for the P wave),

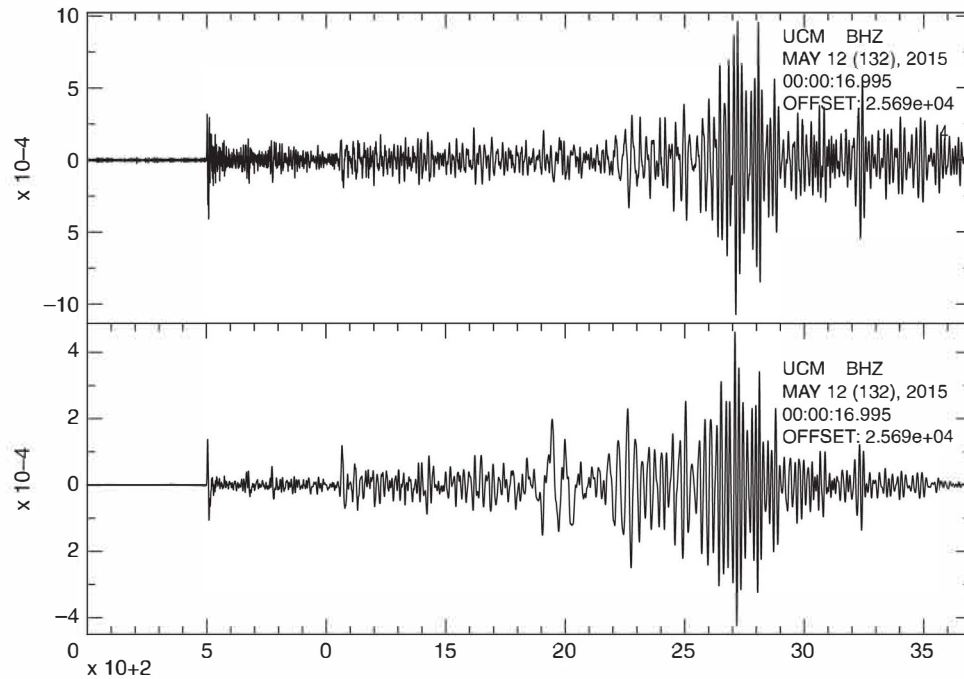


Fig. 3.11

Ground motion displacement of the vertical component of the seismograms of Fig. 3.9. Top: velocity response seismogram. Bottom: ground displacement (10^{-5} m).

for example, of 350 s. In our case the sampling rate is 20 samples/s, so that we have a total of 7000 samples, which is close to 8192 (2^{13}). Because of the requirements of the *fast Fourier transform* (fft) algorithm used in the program, it is important that the selected number of samples is close to a value which is a power of two (2^N). As we did before, we have first to remove the mean and trend of the data and apply a taper. To estimate the FT we use the *fft* command and to plot the spectra the *psp* command. In Fig. 3.12 we have the amplitude and phase spectra for the selected time window of the ground motion displacement. Amplitude spectra are used in many applications of seismology, for example, in the determination of the source mechanism of earthquakes (Chapter 20).

3.6.5 Filtering seismograms

A frequent problem in the analysis of seismograms is the presence of *noise*, spurious signal components not related to the ground motion produced by earthquakes that we want to study. It is, then, important to remove noise with minimum impact to the signal. The origin of this noise may be either within the seismometer system itself, during the digitization process, or due to other causes such as atmospheric phenomena or the sea motion on stations near to the coast. An important source of noise are the so-called *microseisms* of meteorological origin with periods between four and eight seconds. Removing the noise

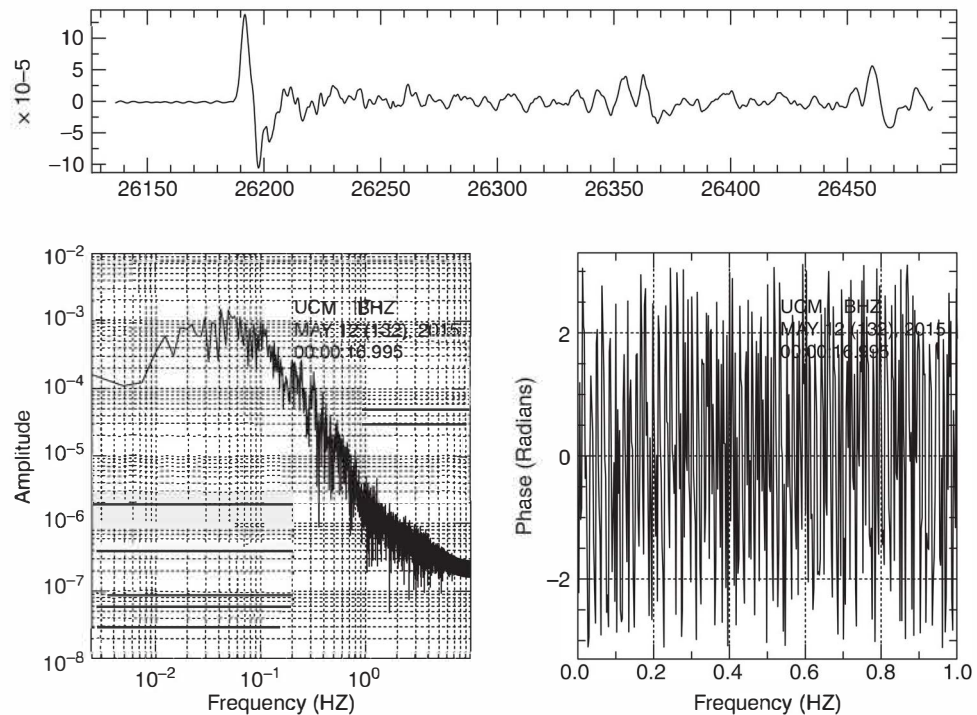


Fig. 3.12 Amplitude and phase spectra of the P wave ground motion displacement of Fig. 3.11.

becomes problematic when the seismic signal and noise have a similar range of frequencies (Scherbaum, p. 134; Zhou, pp. 51–53)

A way to remove the noise in seismograms is by filtering the record: eliminating a selected range of frequencies. To do this we have first to be sure that the frequency of the peak of the noise is different from that of the earthquake signal. For example, if the peak noise frequency is 30 Hz, we can easily remove it because the maximum frequency for an earthquake is about 8–10 Hz. A good practice is to estimate the noise spectra before filtering the seismogram. Figure 3.13 shows the noise amplitude spectra for 350 s before the first arrival of seismic waves for our example, the Nepal earthquake recorded at the UCM station, BHZ component. The peak of the maximum amplitude of the noise is centered at about 0.2 Hz (5 s period). To remove the noise we filter our seismogram using a low-pass filter of three poles with corner frequency 0.3 Hz (3.3 s period). Figure 3.14 shows a time window of 60 s with the arrival of the P wave of the original seismogram and of the filtered seismogram. We can observe that we have removed the high frequencies on the signal, but some noise is still present on the seismogram. To remove it we would have to use a filter with lower corner frequency than 0.2 Hz (period 5 s), but in this case we also remove part of the information of the earthquake signal because its corner frequency is about 0.2–0.3 Hz (3.3–5 s period) (Fig. 3.12). We will study this problem further in Chapter 20.

It is important to point out here that the noise present on seismograms has information about the media through which it is propagating, that is about the Earth's structure. In recent

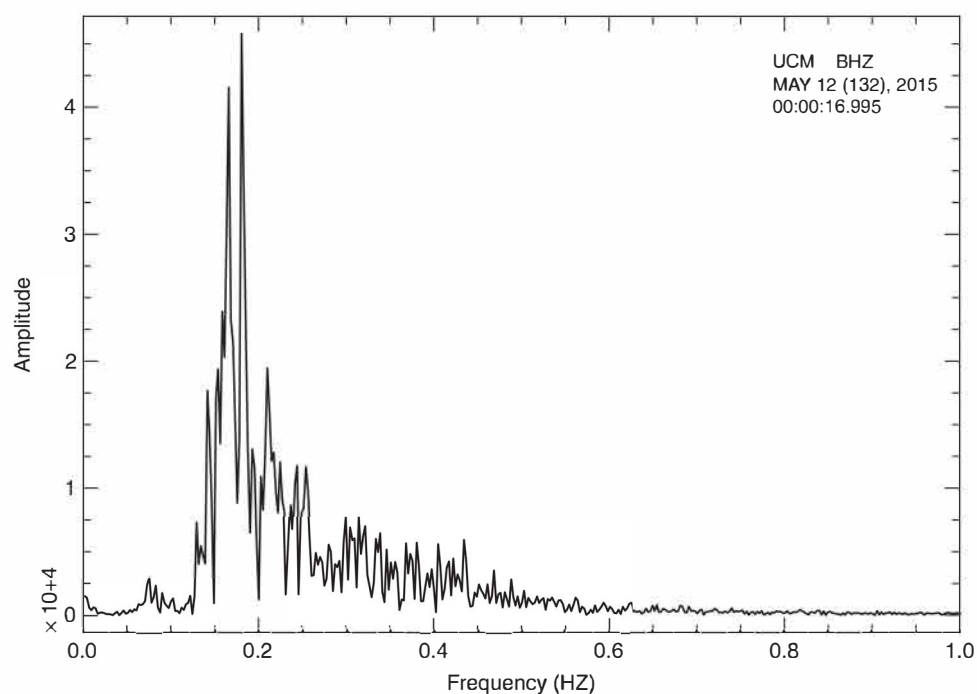


Fig. 3.13 Amplitude spectrum of the noise recorded at the UCM station, BHZ component for the Nepal earthquake. Frequencies are shown in a linear scale.

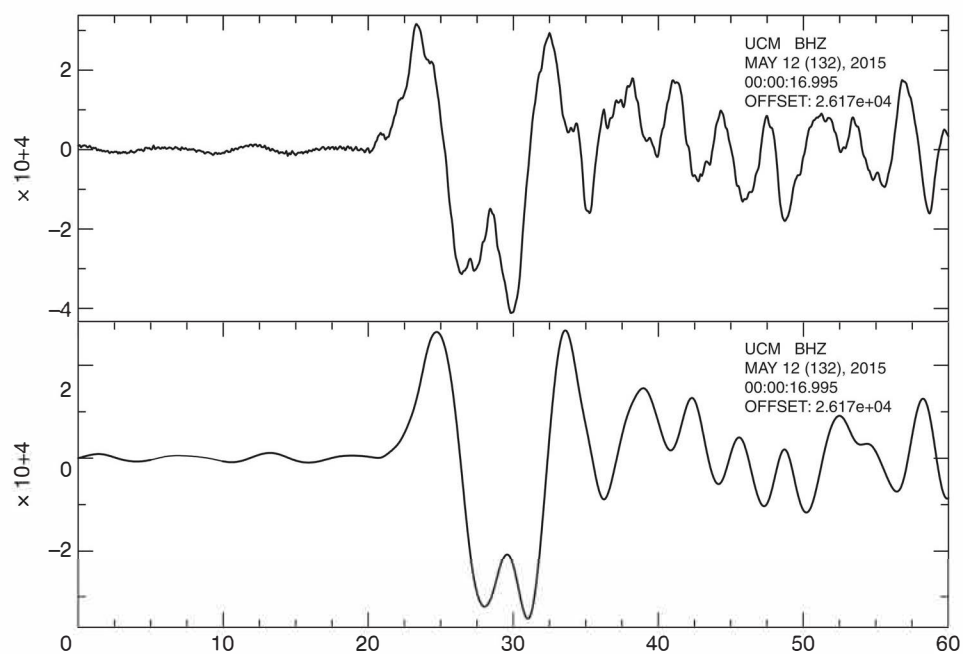


Fig. 3.14 First 60 s of the seismogram recorded at the UCM station, BHZ component for the Nepal earthquake. Top: original seismogram. Bottom: filtered seismogram with a low-pass filter.

years several methods have been developed to study this structure or its site effects on ground motion using environmental or anthropogenic noise ([Chapter 11](#)).

3.7 Accelerographs

Ground motion, which constitutes the observational data for the study of earthquakes can be described in terms of displacement, velocity, and acceleration. Most seismographs are designed to record very weak ground motions from small nearby earthquakes or large distant ones. These instruments are *saturated* and *clipped* with the ground motion of large earthquakes at a near distance making their data unusable ([Fig. 3.8](#)). Moderate to large earthquakes produce near the epicenter, in what is called the *near field*, large ground accelerations at high frequencies that may reach values larger than the acceleration due to gravity (1 g). Seismographs specifically designed to record this type of motion are called *strong motion instruments* or *accelerographs*. In these instruments the relative displacement of the mass corresponds to ground accelerations ([Section 3.3](#)).

Accelerographs were originally very low sensitivity analogical instruments developed for uses in earthquake engineering. Their response is proportional to acceleration, and they record large accelerations up to several times that of gravity. Since the beginning of digital seismology in the late 1970s, accelerographs have also become digital. Digital accelerographs allow the recording of the ground acceleration of large earthquakes at very short distances without saturation. By integration we can find the corresponding ground velocity and displacement ([Fig. 3.15](#)). An important difference between velocity and acceleration records is in the frequency contents; the acceleration record has a broader range of high frequencies in comparison to ground velocity records ([Fig. 3.15](#)).

Integrating accelerograms in order to find velocity and displacement is unfortunately not easy, especially at low frequency, because of problems with rotation of the vertical axis of the instruments and a number of non-linear effects. Although there are many ways to correct ground motion for these instruments, this is still a difficult problem (see Boore *et al.*, 2002, for a detailed discussion). Another problem is that the double integration (from acceleration to displacement) introduces a low frequency noise and it is sometimes necessary to filter the integrated signal.

Accelerometers have a very simple response, due to the fact that the output is linearly proportional to the acceleration and there is no phase shift. Usually BB velocity records have linear bandwidths from 0.01 to 50 Hz and gain from 100 to 5000 V/ms⁻¹ and accelerometers have linear bandwidths from 0 to 100 Hz and gain from 1 to 100 V/g (V/10 ms⁻²) (Havskov and Ottemöller, 2010). Actually, the low frequency pass band of accelerograms is limited by the total length of time of the records, usually a few hundred seconds. For the moment, at least, few accelerograms are recorded continuously. In general, accelerometers have more instrument noise but the large dynamic range is a clear advantage for recording moderate to large earthquakes in the near field without saturation.

In some cases accelerographs give records in ASCII format (normally called *file.raw*) and units of cm/s²; in this case there is no need to remove the instrumental response. In other

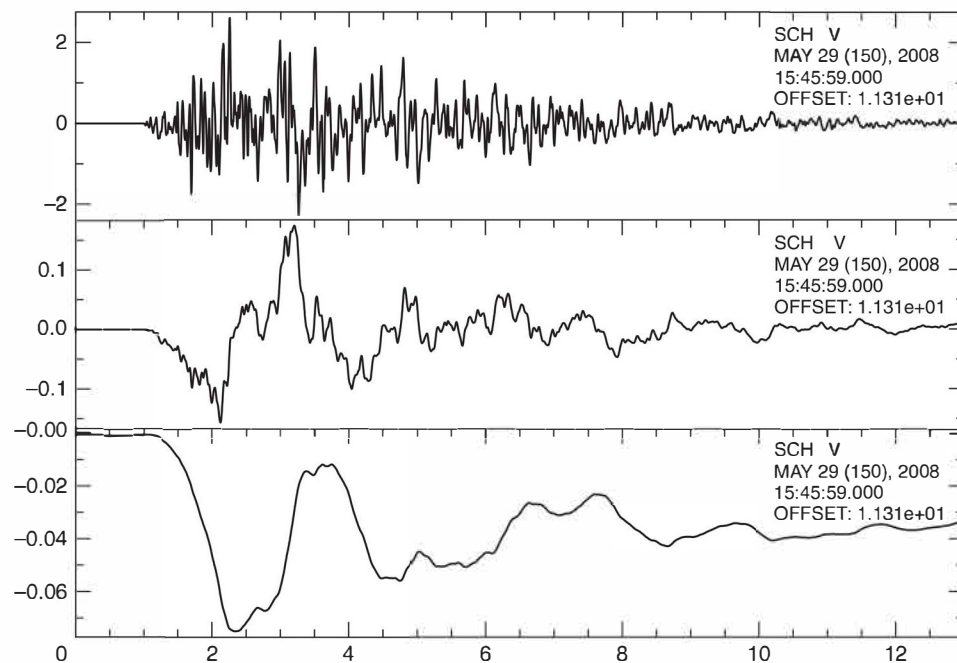


Fig. 3.15 Accelerogram of the station SCH (epicentral distance 6 km) for the Iceland 29 May 2008 earthquake. Records of acceleration, velocity, and displacement are shown (data from Professor Ragnar Sigbjörnsson).

cases, both velocity and acceleration records are given in *SEED* format, which we have to convert to *SAC* with amplitude units in counts. In this case we need to know the instrumental response to compute ground velocity and acceleration.

The observation of ground accelerations in the near field is of great interest in earthquake engineering ([Chapter 21](#)). Antiseismic designs of buildings are based on knowledge of the accelerations expected from nearby earthquakes. Large arrays of accelerographs allow the observation of ground accelerations at different distances and different sites. Thus the attenuation of acceleration in the near field can be measured directly rather than its having to be determined from intensity data. Accelerographs can be located on the ground (free field) or in buildings at various levels. In this way, responses of buildings can be obtained at various heights. The properties of the near-surface layers (rock, soft or hard soil) are important for high-frequency ground motion and constitute the so-called *site effects*. These effects are determined by using strong-motion instruments installed on various kinds of soil.

3.8 Other types of seismologic instruments

In addition to the instruments we have described, there are other seismologic instruments that will be discussed briefly. The *Benioff strain seismometer*, developed in around 1935, consists of a long (10–100 m) quartz bar with one end fixed to a pillar and the other free and

near to a second pillar. By means of a transducer, variations of the distance between the free end of the bar and the pillar fixed to the ground are measured. This distance represents the deformation of the ground between the two pillars. The current from the transducer is passed to a galvanometer and the resulting system has a response similar to that of a long-period seismograph.

A special type of seismograph are portable instruments designed to record small, near earthquakes. They are short-period instruments with high gain that can be installed in the field and operated with batteries. The first models, developed during the 1960s, used continuous analog recording (on smoked paper or by pen and ink). Modern instruments employ digital high dynamic range recording and can be used for continuous recording or as event recorders. The latter are triggered, when the ground motion reaches a certain threshold value, by a triggering algorithm. These instruments are very useful for detailed studies of the seismicity of a limited region, studies of aftershocks after a large earthquake, and monitoring volcanic activity.

Seismographs designed to record earthquakes at the bottom of the sea are known as *ocean bottom seismographs* (OBS). Owing to the fact that oceans cover 70% of the Earth's surface, these instruments are becoming more and more important as a means of improving the geographic distribution of stations. Also, they are very useful for studying the seismicity of oceanic regions. There are several models of OBS; some are connected to a buoy at the surface of the sea with a radio transmitter whereas others record on an internal system and must be recovered after a certain length of time (Fig. 3.16). These types of instrument are fundamental for tsunami early warning systems to give the alert and to mitigate the damage from large offshore earthquakes (see Chapter 21).

Two other types of instrument used in studies related to seismology, which are not seismographs, are tiltmeters and dilatometers. The first measure changes in the tilt of the Earth's surface and the second measure changes in the volume of rocks. Both types are used in connection with prediction programs to measure the accumulation of strain in a region.

A new source of information about ground displacements produced by earthquakes is provided by *global positioning systems* (GPS). GPS is a global navigation satellite system which estimates with high accuracy the position of a point with respect to a terrestrial reference frame. Repeated measurements of the location of the recording site give us its displacements with respect to the reference frame. In this way we can obtain an estimate of the motion of a particular point during a time interval from successive positioning of the recording site. For this purpose two different methodologies are commonly used: *precise point positioning* and *differential positioning* (Larson, 2009). The processing of GPS data is carried out by geodesists and it is not the particular concern of this textbook (Misra and Enge, 2012).

GPS data measured at time intervals of years or months is being used as a standard method to determine the velocity and direction of plate motion in geodynamic studies (Prawirodirdjo and Bock, 2004). For seismological purposes, that is to measure the displacements that take place during an earthquake (the *coseismic* displacements), we need measurements repeated at very short intervals of a few seconds. This has only been possible since the late 1990s and at present the use of high rate GPS sampling (1 Hz and even higher, up to 50 Hz) has opened the way to the use of GPS instruments to measure dynamic ground deformation during large



Fig. 3.16

Deployment of an OBS at the Alboran Sea (between Spain and Morocco) by a ship of the Spanish Navy (courtesy of Dr. A. Pazos, from the Real Observatorio de la Armada, San Fernando). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

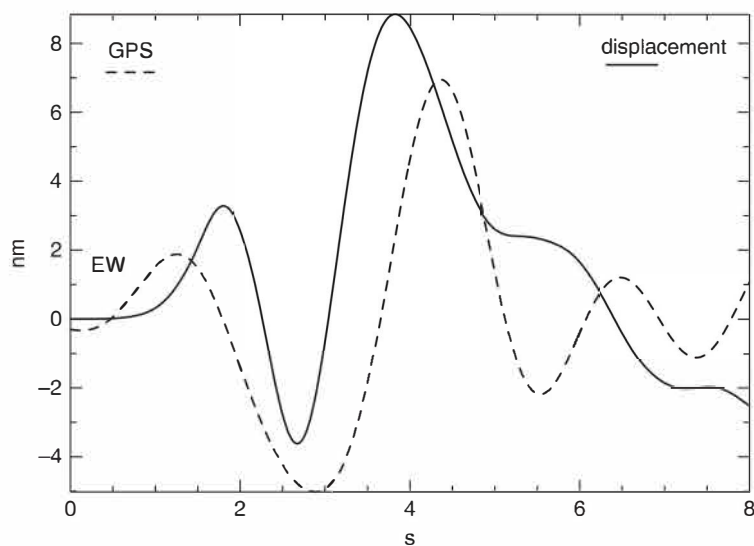


Fig. 3.17

Displacement recorded by GPS instrument and integrated from a strong motion instrument (E–W component) for the 2011 Lorca earthquake ($M_w = 5.1$) recorded at 4 km distance (courtesy of Professor C. Pro).

earthquakes. The recording of such repeated measurements of the position of the instrument site gives what can be considered to be a new type of seismograms (Larson, 2009). GPS data provide a direct measure of the ground displacement, which is an advantage over the velocity response seismographs or accelerographs, where displacements have to be found through integration, with possible introduction of errors. Since present GPS error margins in position are of the order of 1 cm, for seismological purpose only large displacements of the order of a few cm can be recorded. The use of GPS as a seismometer is limited, thus, to large earthquakes or to shallow focus events of lower magnitude at very short distances. Their recordings correspond, generally, to surface waves. When GPS and strong motion instruments are at the same site it is possible to compare the ground displacement measured directly by GPS and the displacement computed by the integration of strong motion acceleration records (Fig. 3.17).

3.9 Summary

The development of seismographs began at the end of the eighteenth century, the first mechanical instruments being based on the motion of a pendulum. In about 1910 electromagnetic seismographs were developed and from 1970 electronic amplification was introduced followed by digital recording, and finally, since 1982, broad-band instruments that are widely used at present.

The theory of the seismograph, for mechanical instruments, expresses the relation of the motion of the ground with that between the pendulum and the frame of the instrument, which is recorded using a magnification system. Thus we can obtain from the records of seismographs the true ground motion. This theory is the basis for that in electromagnetic and digital instruments.

The processing of modern digital data consists of: data acquisition; removing the instrumental response to obtain the ground motion; performance of spectral analysis; and filtering of the signal. In this way we obtain the ground motion in digital form. For close distances accelerographs are used from which we can obtain the ground motion produced by earthquakes at nearby distances. Other types of seismological instruments are OBS for recording at sea, and high rate sampling GPS instruments which provide direct measurement of ground displacement.

3.10 Problems

- P3.1 For a 12 bit digital seismograph for velocity ground motion the output is an electrical signal of 2.5 V. How much is the equivalence between counts and ground motion if the value of the open circuit generator G_o is 342 V/ms^{-1} , the damping resistor $R_x = 4250 \text{ ohms}$, the coil resistance $R_L = 5950 \text{ ohms}$, the gain $A_a = 1976$?

P3.2 The instrument response for a broad-band station (velocity instrument) is given by

#	+	Response (poles & zeros),	+
#	+	+	+
Transfer function type:		B (analog (Hz))	
Response in units lookup:		M/S (velocity in meters per second)	
Response out units lookup:		V (volts)	
A ₀ normalization factor:		1.25314E+11	
Number of zeros:		5	
Number of poles:		12	

Complex zeros:

I	Real	Imaginary
0	0.000000E+00	0.000000E+00
1	0.000000E+00	0.000000E+00
2	-1.250000E+01	0.000000E+00
3	-2.427180E-02	0.000000E+00
4	-2.427180E-02	0.000000E+00

Complex poles:

I	Real	Imaginary
0	-1.968420E-03	1.950130E-03
1	-1.968420E-03	-1.950130E-03
2	-2.260080E-02	0.000000E+00
3	-2.617810E-02	0.000000E+00
4	-6.375250E+00	7.704300E+00
5	-6.375250E+00	-7.704300E+00
6	-4.830000E+01	1.292710E+01
7	-4.830000E+01	-1.292710E+01
8	-3.535540E+01	3.535530E+01
9	-3.535540E+01	-3.535530E+01
10	-1.294000E+01	4.829650E+01
11	-1.294000E+01	-4.829650E+01

Sensitivity: 4.509750E+08

Write the displacement instrument response for this station.

- P3.3 For the digital BHZ seismogram of CART (Cartagena, Spain) station for the earthquake of Mozambique 22 February 2006, obtain the ground motion displacement (material on EM_P3.3).
- P3.4 For the horizontal components LHE and LHN of station MELI (Melilla, North Africa) for the Peru earthquake of 23 June 2001 ($M_w = 8.3$), obtain the radial and transverse components (material on EM_P3.4).
- P3.5 For the seismogram of Problem P3.3 obtain the amplitude spectrum for the P wave.

Ground motions produced by earthquakes and recorded by seismographs, as explained in [Chapter 3](#), constitute the basic observations for the quantitative study of earthquakes. In order to formalize and quantify the information about the motion produced by earthquakes we use the principles and equations of the mechanics of *continuous media*. For this purpose, we approximate the Earth by a continuous deformable medium to which the equations of mechanics can be applied. We start by recalling some of the basic ideas on the mechanical behavior of continuous media. These ideas constitute the fundamentals necessary to understand the motion produced by earthquakes. It is, then, very important to understand these in order to quantify and formalize our knowledge. First we explain the basic concepts of displacement, strain, stress, and elasticity coefficients in order to proceed to presentation of the fundamental equations of motion, with consideration of energy and heat release.

4.1 Displacement, strain, and stress

A continuous medium is an idealization of a material in which the distance between two contiguous points can be made infinitesimally small. In this idealization, the granular structure of the materials of the Earth and their molecular and atomic nature are not considered. In a continuous medium the term particle is used to mean a geometric point without dimensions. Density and mechanical properties of continuous media are considered to be continuous functions of the spatial coordinates and time. Motion of points in a continuous medium can be described by what are known as *Lagrangian* and *Eulerian* descriptions (after Leonhard Euler and Joseph Louis Lagrange). In seismology we use the linearized theory of elasticity adopting the Lagrangian viewpoint, which considers points in the medium as individual moving particles (Dahlen and Tromp, 1998, 26–36).

The study of the mechanical behavior of continuous media can be traced back to the discovery in 1660 by Robert Hooke of a linear law that relates applied forces and deformations in an elastic body and its applications formulated by Edmé Mariotte at around 1680. The first studies on the behavior of elastic materials were those of Daniel Bernoulli, Euler, Lagrange, and Charles Coulomb. In 1827, Claude-Louis Navier established the general equations for equilibrium and vibration in elastic solids. This work was continued by Augustine Louis Cauchy, Siméon-Denis Poisson, Gabriel Lamé, Gustav Robert Kirchhoff, William Thomson (Lord Kelvin), John William Strutt (Lord Rayleigh), Augustus E. H. Love, and Horace Lamb, among others.

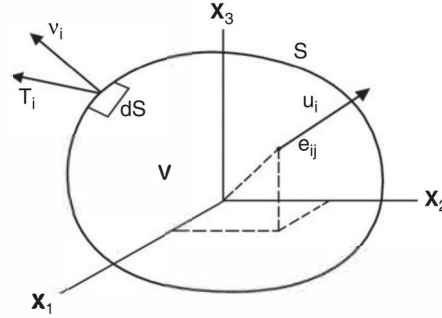


Fig. 4.1 The displacements (u_i), strain (e_{ij}), and stress $\mathbf{T}$, in an elastic medium.

Let us consider, in a continuous medium, a region of volume V surrounded by a closed surface S , in such a way that there is matter on both sides of the surface. For each point inside V we can define *displacements*, *strains*, and *stresses*, as continuous functions of the spatial coordinates and time (Fig. 4.1). *Displacements* at a point in a continuous medium are given by the vector $\mathbf{u}$, *velocities* and *accelerations* by their time derivatives: $d\mathbf{u}/dt$ and $d^2\mathbf{u}/dt^2$. The ground displacements together with their velocities and accelerations recorded by seismographs (Chapter 3) constitute the basic quantitative observations of the motion produced by earthquakes.

If l is the distance between two points, for a rigid-body displacement $\mathbf{u}$ is the same for all points and l remains constant. When a body is deformed, $\mathbf{u}$ is different for different points and the distance between two points changes from l to l' . The *strain* is defined as the change of this distance per unit distance: $e = (l' - l)/l$ (a non-dimensional quantity). In the simplest one-dimensional case,

$$l = dx,$$

and

$$l' = dx + \frac{\partial u}{\partial x} dx.$$

And the strain is given by

$$e = \frac{l' - l}{l} = \frac{\partial u}{\partial x}. \quad (4.1)$$

The *deformation* or *strain* in a body depends on the derivatives of the displacements. Because the strain implies variations in the displacements in the directions of the three spatial coordinates, in the general case, it must be represented by a tensor (for vectors and tensors see Appendix 1). In the case of infinitesimal deformations, the strain can be represented by the *Cauchy tensor* $\mathbf{e}$, which, if $\mathbf{u}$ varies continuously and slowly with position, is given by

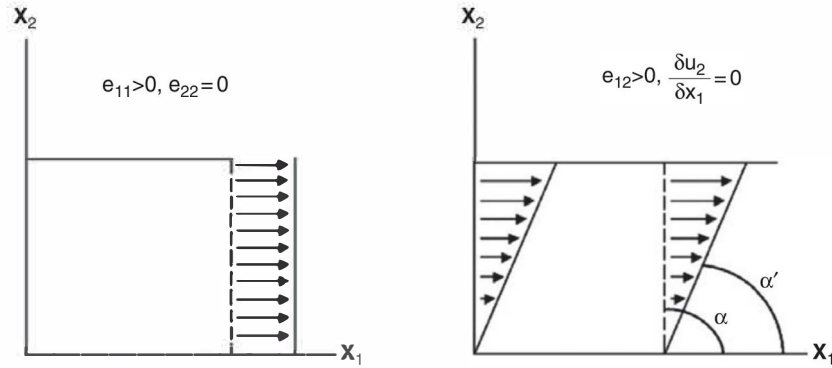


Fig. 4.2

Representations of the longitudinal strain e_{11} and shear strain e_{12} .

$$e_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \frac{1}{2} (u_{i,j} + u_{j,i}). \quad (4.2)$$

In this expression we have used the index notation and commas to represent partial derivatives with respect to the corresponding spatial coordinate ([Appendix 1](#)). This definition implies that the increments in the displacements ($\delta \mathbf{u}$) are smaller than those in position ($\delta \mathbf{x}$). The tensor $\mathbf{e}$ is, by definition, symmetric ($e_{ij} = e_{ji}$) and only six of its nine components are different. The components of the strain tensor $\mathbf{e}$ have a simple geometric interpretation. Those with indexes $i = j$ correspond to *longitudinal deformations* along the axes of coordinates. For $i \neq j$, they correspond to variations in the angles between the directions of the axes of the deformed and un-deformed states; deformations of this type are called *shear strains* ([Fig. 4.2](#)).

Another important tensor related to changes in the displacements in a deformable continuous body is the *rotation tensor*,

$$\omega_{ij} = \frac{1}{2} (u_{i,j} - u_{j,i}). \quad (4.3)$$

The rotation tensor ω is antisymmetric. Using the tensors $\mathbf{e}$ and ω the partial derivatives of the displacements are given by

$$u_{i,j} = e_{ij} + \omega_{ij}. \quad (4.4)$$

The rotation tensor ω_{ij} is related to the curl of the displacements,

$$\begin{aligned} \omega &= \nabla \times \mathbf{u} \\ \omega_i &= e_{ijk} u_{k,j} \end{aligned} \quad (4.5)$$

by the equation

$$\omega_i = 2e_{ijk}\omega_{jk}, \quad (4.6)$$

where e_{ijk} is the *alternating* or *permutation tensor* ([Appendix 1](#)). According to (4.4), the partial derivatives of $\mathbf{u}$ are totally defined by the two tensors e_{ij} and ω_{ij} and contain all the

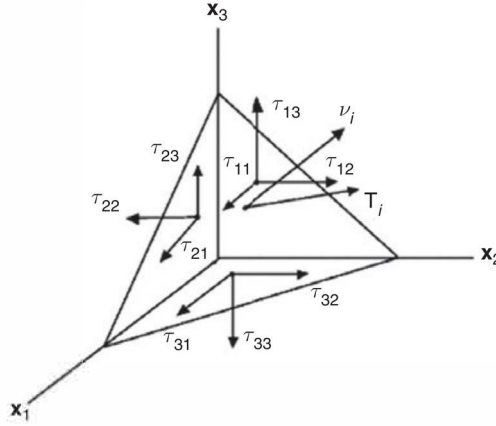


Fig. 4.3

The components of the stress tensor, τ_{ij} , through three orthogonal planes and of the stress vector T_i through a plane normal to the unit vector v_i .

information about the deformation of the body. This means that the variations of displacements from one point to another in a deformable medium include both strains and rotations.

Stresses at a point inside V are defined as the limit of the quotients of the forces that act at this point per unit surface through a plane with a certain orientation. Both the orientation of the force $\mathbf{F}$ and that of the plane ΔS are included in the definition of the stress. The stress through a plane with normal $\mathbf{v}$ is represented by a vector $\mathbf{T}$,

$$\mathbf{T} = \lim_{\Delta S \rightarrow 0} \frac{\mathbf{F}}{\Delta S}. \quad (4.7)$$

Stress exists only if there is material on both sides of the surface. Stress can also be represented by a second-order tensor $\boldsymbol{\tau}$ with nine elements (τ_{ij}) that are the stresses through three orthogonal planes. In Cartesian coordinates, these planes are normal to the three axes (x_1 , x_2 , and x_3) (Fig. 4.3). With this tensor we can define the state of stress at a point through a plane of arbitrary orientation. For a plane with a normal given by the unit vector $\mathbf{v}$, Cauchy's relation between the vector $\mathbf{T}$ and the tensor $\boldsymbol{\tau}$ is

$$T_i = \sum_{j=1}^3 \tau_{ij} v_j = \tau_{ij} v_j. \quad (4.8)$$

In this equation we have used the convention that repeated indexes in the same term are summed through their three values (Albert Einstein summation convention), a convention that will be used in all following equations. Vector $\mathbf{T}$ needs to add the information about the orientation of the plane on which it is acting (its normal $\mathbf{v}$), whereas $\boldsymbol{\tau}$ has all the information about the state of stress at a point of the medium. It is easy to show that, in the absence of external moments, the tensor $\boldsymbol{\tau}$ is symmetric ($\tau_{ij} = \tau_{ji}$) and only six of the nine components are different. With respect to the three planes normal to the coordinate axes, the

components of τ_{ij} with indexes $i = j$ correspond to *normal stresses* and those with $i \neq j$ correspond to tangential or *shear stresses*.

4.1.1 Eigenvalues and eigenvectors

The tensors τ_{ij} and e_{ij} can be studied in terms of their *eigenvalues* and *eigenvectors*. Since both tensors are represented by 3×3 symmetric matrices, in each case, their three eigenvalues are real and their three eigenvectors mutually orthogonal. The eigenvectors form a system of orthogonal axes such that, when referred to it, all components of the tensor are zero, except those in the main diagonal, where they are the eigenvalues. The reference system formed by the eigenvectors represents the system of *principal axes* of stress and strain. The eigenvalues are also called the *principal values* of the stress and strain. Eigenvectors (v_i) and eigenvalues (σ) are found through the equation

$$(\tau_{ij} - \sigma\delta_{ij})v_i = 0. \quad (4.9)$$

In this equation δ_{ij} is the *Kronecker delta tensor* (after Leopold Kronecker) ([Appendix 1](#)). The eigenvalues are the three roots of the cubic equation resulting from equating to zero the determinant of (4.9),

$$\text{Det}[\tau_{ij} - \sigma\delta_{ij}] = 0. \quad (4.10)$$

The three eigenvectors, v_i^1 , v_i^2 , and v_i^3 , correspond to the three eigenvalues σ_1 , σ_2 , and σ_3 and are obtained by substituting each of these values into equation (4.9). Thus, referred to its principal axes, the tensor τ_{ij} has the following form:

$$\tau_{ij} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}. \quad (4.11)$$

In this system of axes, the tensor τ_{ij} has only normal components. Regarding the planes normal to the axes, there exist only normal stresses (τ_{ij} , $i = j$) and the shear stresses (τ_{ij} , $i \neq j$) are null. The principal values of the stress, σ_1 , σ_2 , and σ_3 , are, usually, ordered so that σ_1 is the largest and σ_3 the smallest.

A similar result is obtained for the strain tensor e_{ij} . When it is referred to its principal axes, the shear components are null and the tensor takes the form

$$e_{ij} = \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{bmatrix}. \quad (4.12)$$

In both cases, for τ and e , the sum of the elements of the main diagonal is the first invariant of the matrix, so that, for any orientation of the axes,

$$\begin{aligned} \tau_{11} + \tau_{22} + \tau_{33} &= \sigma_1 + \sigma_2 + \sigma_3 = \text{constant} \\ e_{11} + e_{22} + e_{33} &= \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \text{constant}. \end{aligned}$$

In the case of the strain tensor, this sum represents the change in volume per unit volume and it is called the *cubic dilation* θ . It can be obtained from (4.2) that the cubic dilation is equal to the divergence of the displacements,

$$\begin{aligned} e_{11} + e_{22} + e_{33} &= u_{1,1} + u_{2,2} + u_{3,3} = \theta \\ \nabla \cdot \mathbf{u} &= \theta. \end{aligned} \quad (4.13)$$

The tensors τ_{ij} and e_{ij} can be expressed as the sum of two tensors, one isotropic and the other deviatoric:

$$\begin{aligned} \tau_{ij} &= \sigma_0 \delta_{ij} + \tau'_{ij} \\ e_{ij} &= \varepsilon_0 \delta_{ij} + e'_{ij}, \end{aligned} \quad (4.14)$$

where σ_0 and ε_0 are one-third the sums of the principal stresses and strains, respectively:

$$\begin{aligned} \sigma_0 &= \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) \\ \varepsilon_0 &= \frac{1}{3}(\varepsilon_1 + \varepsilon_2 + \varepsilon_3). \end{aligned} \quad (4.15)$$

The *deviatoric stress* τ'_{ij} and *strain* e'_{ij} are, thus, defined by equations (4.14). In the case that the deformation implies only changes in volume without changes in form, $e'_{ij} = 0$. If there are only changes in form, without changes in volume, $\varepsilon_0 = 0$. In this case the strain is purely deviatoric. These concepts will be used in Chapter 17 when we study the source mechanism of earthquakes.

4.2 Elasticity coefficients

The mechanical behavior of a continuous material is defined by the relation between the stress and the strain. For a *linear elastic body*, this relation is given by Hooke's law, which states that the strain is proportional to the stress. Cauchy's formulation in tensor form of this law is

$$\tau_{ij} = C_{ijkl} e_{kl}. \quad (4.16)$$

This equation states the linear relation between the strain and stress tensors and is the foundation of the *theory of linear elasticity*. The fourth-order tensor C_{ijkl} is the tensor of the *elasticity coefficients* or *moduli* which has, in general, 81 component functions of the coordinates. Owing to the symmetry of τ_{ij} and e_{ij} , only 36 are different. For perfect elasticity, there exists a strain-energy function and it follows that $C_{ijkl} = C_{klij}$ and the number of elasticity coefficients is further reduced to 21 (Malvern, 1969). Equation (4.16) can also be expressed in terms of the derivatives of the displacements, by substituting (4.2) into it:

$$\tau_{ij} = C_{ijkl} u_{k,l}. \quad (4.17)$$

The simplest case for the elasticity coefficients corresponds to an *isotropic medium*, that is, a medium with the same properties in all directions. For such a medium, all components of C_{ijkl} can be expressed by two coefficients, λ and μ , called *Lamé's coefficients*:

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}). \quad (4.18)$$

The 21 different components of C_{ijkl} in terms of λ and μ are

$$\begin{aligned} C_{1111} &= C_{2222} = C_{3333} = \lambda + 2\mu \\ C_{1122} &= C_{1133} = C_{2233} = C_{2211} = C_{3311} = C_{3322} = \lambda \\ C_{1212} &= C_{2121} = C_{1221} = C_{2112} = C_{1313} = C_{3131} = C_{1331} = C_{3113} = C_{2323} \\ &= C_{3232} = C_{2332} = C_{3223} = \mu. \end{aligned}$$

By substitution of (4.18) into (4.16), the relation between the stress and the strain for an isotropic medium is found to be

$$\tau_{ij} = \lambda \delta_{ij} e_{kk} + 2\mu e_{ij}. \quad (4.19)$$

In terms of the derivatives of the displacements,

$$\tau_{ij} = \lambda \delta_{ij} u_{k,k} + \mu (u_{i,j} + u_{j,i}). \quad (4.20)$$

From a different point of view, the mechanical behavior of an isotropic elastic medium can be stated in terms of two coefficients: K , the *bulk modulus* that relates the changes in volume without changes in form; and G , the *shear modulus* that relates the changes in form without changes in volume to the stresses that produce them. For the first case,

$$\tau_{11} + \tau_{22} + \tau_{33} = 3K(e_{11} + e_{22} + e_{33}). \quad (4.21)$$

In the case of hydrostatic pressure, the normal stresses are equal ($\tau_{11} = \tau_{22} = \tau_{33} = -P$) and equation (4.20), taking into account (4.13), results in

$$P = -K\theta.$$

The coefficient K represents the quotient relating the pressure to the change in volume it produces and for this reason it is called the bulk modulus,

$$K = -\frac{P}{\theta} = -V \frac{dP}{dV}. \quad (4.22)$$

The second elasticity coefficient G relates changes in form without changes in volume to the shear stresses; the relation between the deviatoric strain and stress is

$$\tau'_{ij} = 2G e'_{ij}. \quad (4.23)$$

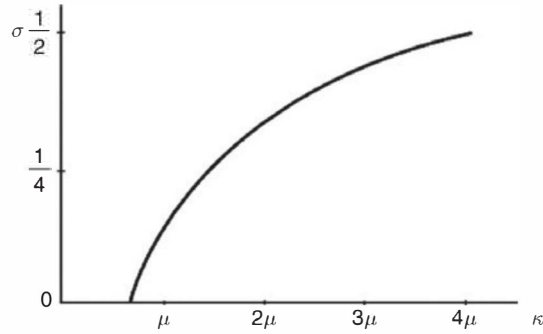
The coefficient G is equivalent to the Lamé coefficient μ , which is also called the *shear* or *rigidity modulus*. By substituting equation (4.21) and (4.23) into (4.19), and considering (4.14), we obtain for $i \neq j$, $G = \mu$, and for $i = j$,

$$K = \lambda + \frac{2}{3}\mu. \quad (4.24)$$

Another elasticity coefficient is *Young's modulus* E , which relates the longitudinal stress and strain in the same direction:

Table 4.1 Correspondence among values of K , λ , and μ for some values of σ .

σ	λ	K
0	0	$\frac{2}{3}\mu$
$\frac{1}{8}$	$\frac{1}{3}\mu$	M
$\frac{1}{4}$	M	$\frac{5}{3}\mu$
$\frac{1}{2}$	K	A

**Fig. 4.4**

The relation between Poisson's ratio σ and the bulk modulus K in terms of the rigidity μ or shear modulus.

$$E = \tau_{11}/e_{11}. \quad (4.25)$$

The relations of E to λ , μ , and K are

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} = \frac{9K\mu}{3K + \mu}. \quad (4.26)$$

In a medium subject only to a longitudinal stress τ_{11} in the direction of the x_1 axis, the quotient relating the strain in a perpendicular direction e_{22} and that in the same direction as the stress e_{11} is called *Poisson's ratio*:

$$\sigma = -e_{22}/e_{11}. \quad (4.27)$$

On substituting this into equation (4.19) and considering that $e_{33} = e_{22}$, the following relation is obtained:

$$\sigma = \frac{\lambda}{2(\lambda + \mu)}. \quad (4.28)$$

Poisson's ratio σ has values between 0 and $\frac{1}{2}$. The correspondence among the values of K , λ , and μ for some values of σ are given in Table 4.1 and shown in Fig. 4.4.

For $\sigma < \frac{1}{8}$, the material changes volume with greater facility than it changes its form ($K < \mu$). For $\sigma > \frac{1}{8}$, we have the opposite case; the material changes form more easily than it changes in volume ($K > \mu$). When $\mu = 0$, the material cannot support any shear stress, which is the case for a fluid. If $\mu = 0$ and $K = \infty$, the fluid is incompressible. If $K = \infty$ and $\mu \neq 0$, the material is an incompressible solid. The condition $\sigma = \frac{1}{3}$, for which $\lambda = \mu$, is known as *Poisson's condition* and the elastic behavior of the material (*Poisson's solid*) is defined by only one parameter. This condition is approximately satisfied for most materials in the Earth's interior (σ varies in the range 0.22–0.35). Solutions of many seismologic problems are greatly simplified by using this approximation.

The units used nowadays in seismology are preferentially those of the International System (IS), although sometimes cgs units are still used. For displacements the units are m (IS) and cm (cgs), for stress Nm^{-2} or Pa (IS) and dyn cm^{-2} (cgs) and also bars (1 bar = 10^6 dyn cm^{-2}) ($10^6 \text{ Pa} = 1 \text{ MPa} = 10 \text{ bars}$). The elastic coefficients K , λ , μ , and E have the same units as stress. Strain is adimensional or with units of mm^{-1} (IS) and cm cm^{-1} (cgs).

4.3 The influence of temperature

In the process of deformation of a continuous elastic medium, temperature changes and heat generation and flow may be produced. For example, in the generation of earthquakes by faulting part of the energy release is transformed into heat. This means that the thermodynamic conditions must be considered. We will give some very basic considerations about this problem. In the first place, if during a deformation process the temperature changes from an initial value T_0 to another T , the general relation between stress and strain is given by the Duhamel–Neumann formulation of Hooke's law (named after Jean M.C. Duhamel and Franz Ernst Neumann),

$$\tau_{ij} = C_{ijkl}e_{kl} - \beta_{ij}(T - T_0), \quad (4.29)$$

where β_{ij} are the *coefficients of linear thermal expansion*. For an isotropic medium the relation corresponding to (4.19) is

$$\tau_{ij} = \delta_{ij}\lambda e_{kk} + 2\mu e_{ij} - K\alpha(T - T_0)\delta_{ij}, \quad (4.30)$$

where, in the last term, α is the scalar *coefficient of thermal expansion*. If there are no applied stresses ($\tau_{ij} = 0$), then, considering equations (4.21) and (4.24), we obtain from (4.30),

$$e_{kk} = \theta = \alpha(T - T_0). \quad (4.31)$$

If the temperature increases ($T > T_0$), θ is positive and there is an increase in volume. The coefficient α relates the increase in volume to the increase in temperature at constant pressure:

$$\alpha = \left(\frac{\partial V}{\partial T} \right)_P. \quad (4.32)$$

We have defined K as the quotient relating the applied pressure to changes in volume; its inverse $1/K$ represents the change in volume with a change in pressure. Since, in the definition of K , we have not considered changes in temperature, this coefficient is valid for processes with a constant temperature and is called the *isothermal bulk modulus*:

$$\frac{1}{K} = \left(\frac{\partial V}{\partial P} \right)_T. \quad (4.33)$$

In processes with changes in temperature when there is no heat transfer, that is, adiabatic processes, we can define the *adiabatic bulk modulus* K_a . Its inverse $1/K_a$ represents the change in volume with a change in pressure at constant entropy:

$$\frac{1}{K_a} = \left(\frac{\partial V}{\partial P} \right)_S. \quad (4.34)$$

From the basic thermodynamic equations, the increments of heat dQ and enthalpy dH are

$$dQ = \left(\frac{\partial Q}{\partial T} \right)_P dT + \left(\frac{\partial Q}{\partial P} \right)_T dP, \quad (4.35)$$

$$dH = T dS + V dP. \quad (4.36)$$

From these two equations and thermodynamic relations, we find the relation between the change of volume with pressure at constant entropy and at constant pressure related with the change of volume with temperature,

$$\left(\frac{\partial V}{\partial P} \right)_S = \left(\frac{\partial V}{\partial P} \right)_T + \frac{T}{C_P} \left(\frac{\partial V}{\partial T} \right)_P^2. \quad (4.37)$$

If we substitute (4.32), (4.33), and (4.34) into (4.37), we obtain

$$\frac{1}{K_a} = \frac{1}{K} - \frac{T}{\alpha^2 C_P}, \quad (4.38)$$

where

$$C_P = \left(\frac{\partial Q}{\partial T} \right)_P \quad (4.39)$$

(C_P is the *specific heat* at constant pressure). Equation (4.38) relates isothermal and adiabatic bulk moduli.

In conclusion, the coefficients that define the mechanical behavior of isotropic elastic media are K , μ , and ρ . If there are changes in temperature we have to add α and, for adiabatic processes, substitute K for K_a .

4.4 Work, energy, and heat considerations

In an elastic medium, stresses that produce deformations result in an amount of *work* dW . The work done by each increment of strain per unit volume is

$$dW = \tau_{ij} de_{ij}. \quad (4.40)$$

If the medium is perfectly elastic there is no dissipation of *energy* and dW is an exact differential that depends uniquely on the strain:

$$dW = \frac{\partial W}{\partial e_{ij}} de_{ij}. \quad (4.41)$$

From (4.40) and (4.41) we can derive

$$\tau_{ij} = \frac{\partial W}{\partial e_{ij}}. \quad (4.42)$$

In this expression W can be considered as the potential function of elastic energy. If we replace τ_{ij} in (4.40) by its value in terms of e_{ij} (4.16), we obtain

$$dW = C_{ijkl} e_{kl} de_{ij}. \quad (4.43)$$

If we integrate this expression, the energy per unit volume for an elastic body is given by

$$W = \frac{1}{2} C_{ijkl} e_{kl} e_{ij} = \frac{1}{2} \tau_{ij} e_{ij}, \quad (4.44)$$

and for an isotropic body,

$$W = \frac{1}{2} \lambda (e_{11} + e_{22} + e_{33})^2 + \mu e_{ij} e_{ij}. \quad (4.45)$$

Let us now consider the thermodynamic problem when part of the work is dissipated as heat. The change in *internal energy* dE in an elastic body subject to changes in work dW and heat dQ is given by

$$dE = dW + dQ.$$

On substituting for dW according to (4.40) and dQ in terms of the change in *entropy* dS ($dS = dQ/T$), we have that

$$dE = \tau_{ij} de_{ij} + T dS. \quad (4.46)$$

In an *adiabatic* process with reversible infinitesimal deformations, there is no generation or transport of heat and the entropy is constant, so

$$dE = \tau_{ij} de_{ij}. \quad (4.47)$$

This equation can also be written as

$$\left(\frac{\partial E}{\partial e_{ij}}\right)_S = \tau_{ij}. \quad (4.48)$$

On comparing it with equation (4.42), this expression shows that, in an elastic body, under adiabatic conditions, the internal energy E is formed only by the *elastic potential energy* W .

In order to consider *isothermal* processes with reversible heat conduction, we introduce the *free energy* or *Helmholtz function* A (named after Hermann Helmholtz) defined in the form

$$A = E - TS. \quad (4.49)$$

If we take differentials of this equation and substitute for dE from (4.47), since the temperature and entropy are constant, we obtain

$$dA = \tau_{ij} de_{ij}.$$

This expression can also be written as

$$\left(\frac{\partial A}{\partial e_{ij}}\right)_T = \tau_{ij}. \quad (4.50)$$

For isothermal processes, the elastic energy potential function W is given by the free energy A . In conclusion, in a perfectly elastic body there exists an elastic energy potential function of the strain (W) from which we can derive the stress. This function is given by the internal energy E in adiabatic processes (4.48) and by the free energy A in isothermal processes (4.50).

If there is production of heat in the volume V and flux of heat through its surface S , the time rate of the *internal energy density* η (internal energy E per unit mass) can be written as

$$\int_V \rho \frac{\partial \eta}{\partial t} dV = \int_V \tau_{ij} \frac{\partial e_{ij}}{\partial t} dV + \int_V \rho \varepsilon dV - \int_S q_i v_i dS, \quad (4.51)$$

where the first term on the right-hand side is the energy rate produced in the process of deformation (τ_{ij} are the applied stresses), the second corresponds to the heat production inside V (ε is the heat produced per unit mass and unit time), and the third term corresponds to the heat flux through the surface S (q_i is the heat flux per unit surface and unit time, positive for outflow and negative for inflow, and v_i is the unit vector normal to the surface element dS). Transforming the surface integral into a volume integral for q_i using the *divergence theorem* or *Gauss theorem* (first derived by Carl Friedrich Gauss, this theorem will be used frequently to pass from surface to volume integrals),

$$\int_S F_i(x_j) dS = \int_V \frac{\partial F_i(x_j)}{\partial x_i} dV, \quad (4.52)$$

and extending the medium to the infinite space, we obtain equation (4.51) in differential form,

$$\rho \frac{\partial \eta}{\partial t} = \tau_{ij} \frac{\partial e_{ij}}{\partial t} + \rho \varepsilon - \frac{\partial q_i}{\partial x_i}. \quad (4.53)$$

The rate of internal energy depends on the change in time of the deformations plus the heat production and minus the heat flux. If the process is not reversible (physical processes are not reversible) then there is an increase in entropy. This can be expressed in terms of the specific entropy s (entropy per unit mass) in the form,

$$\frac{d}{dt} \int_V \rho s dV \geq \int_V \frac{\rho \varepsilon}{T} dV - \int_S \frac{q_i v_i}{T} dS, \quad (4.54)$$

and in differential form,

$$\rho \frac{ds}{dt} \geq \frac{\rho \varepsilon}{T} - \frac{\partial}{\partial x_i} \left(\frac{q_i}{T} \right). \quad (4.55)$$

The increase in entropy depends on heat production and flux. This is the *Clausius–Duhem inequality* for all physical irreversible processes, named after Rudolf Clausius and Pierre Duhem, that becomes an equality only for reversible processes.

These fundamental concepts and relations are very important in order to understand the motion produced by earthquakes as we approximate the Earth as an elastic body, in which by tectonic processes, displacements, strains, and stresses are produced with changes in internal energy and heat conditions.

4.5 Equations of continuity and motion

The fundamental equations that rule the mechanical behavior of an elastic medium are those of continuity and motion. The first is a consequence of the principle of conservation of mass and energy and the second is a consequence of Newton's second law. The application of these equations to the processes in the Earth derived from the occurrence of earthquakes constitutes the fundamentals of theoretical seismology.

4.5.1 The equations of continuity of mass and energy

The mass contained in a volume V of a continuous medium of density $\rho(\mathbf{x}, t)$, a function of the spatial coordinates and time, is given by

$$M(t) = \int_V \rho(\mathbf{x}, t) dV. \quad (4.56)$$

The principle of *conservation of mass* states that mass is conserved in all physical processes. This may be expressed mathematically using the concept of the *material derivative* with respect to time (D/Dt) which must be zero. For a volume V , contained inside a closed surface S , the material derivative of the mass is (Malvern, 1969)

$$\frac{D}{Dt} \int_V \rho dV = \int_V \frac{\partial \rho}{\partial t} dV + \int_S \rho v_j v_j dS. \quad (4.57)$$

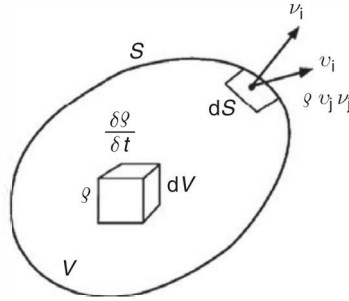


Fig. 4.5

The change with time of the mass inside volume V and the flux of mass through its external surface S .

On the right-hand side of (4.57), the first integral represents the change of mass with time inside the volume V and the second the flux of mass through the surface S . For each element of the surface dS with normal $\mathbf{n}$, the velocity of mass flow is $\mathbf{v}$ and $\mathbf{H} = \rho\mathbf{v}$ the flux per unit surface area (Fig. 4.5). The conservation of mass implies that expression (4.57) is null. Therefore the variation of mass with time inside the volume V must equal the flux through its surface S . If we apply Gauss's theorem (4.52) to convert the surface integral into a volume integral in equation (4.57) and equate it to zero, we obtain

$$\int_V \left(\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} (\rho v_i) \right) dV = 0. \quad (4.58)$$

In differential form this expression is

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} (\rho v_i) = 0. \quad (4.59)$$

In an analogous way, if $w(\mathbf{x}, t)$ is the energy density per unit volume and the flux of energy $\mathbf{H} = w\mathbf{U}$, then the *conservation of energy* in a volume V is given by

$$\int_V \left(\frac{\partial w}{\partial t} + \frac{\partial}{\partial x_i} (w U_i) \right) dV = 0, \quad (4.60)$$

where now $\mathbf{U}$ is the velocity of the energy flow through the surface S that surrounds V . In differential form the resulting equation is similar to (4.59). The equations of continuity (4.58) and (4.60) show the relation between the change with time of the mass and energy inside a volume V and their flow through the surface S which surrounds it. If the change is negative (the mass or energy in V diminishes), the flow is positive (going out through S), and vice versa if the change is positive.

4.5.2 The equation of motion or momentum

The motion at each point inside a volume V is determined by the forces acting in its interior and stresses on its external surface. Their relation is given by the fundamental Newton's *second law of motion* that states that the sum of the forces and stresses equals

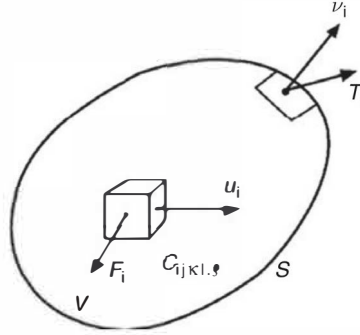


Fig. 4.6

The displacement $\mathbf{u}$, force per unit volume $\mathbf{F}$ (the body force), and stress per unit surface area $\mathbf{T}$, acting on an elastic medium of density ρ and elastic coefficients C_{ijkl} .

the time derivative of the linear momentum. For a continuous body, according to Euler's formulation, this law is given by

$$\int_V F_i dV + \int_S T_i dS = \frac{d}{dt} \int_V \rho v_i dV, \quad (4.61)$$

where $\mathbf{F}$ are forces acting on volume elements dV , called *body forces*, $\mathbf{T}$ stresses acting on surface elements dS , and $\mathbf{v}$ is the velocity at each point of the volume (Fig. 4.6). On substituting for the vector $\mathbf{T}$ the tensor $\boldsymbol{\tau}$, according to (4.8), and applying Gauss's theorem (4.52) to the surface integral to convert it into a volume integral, we obtain

$$\int_V \left(F_i + \frac{\partial \tau_{ij}}{\partial x_j} \right) dV = \frac{d}{dt} \int_V \rho v_i dV. \quad (4.62)$$

If the density is constant with time, this equation becomes in differential form,

$$\begin{aligned} \frac{\partial \tau_{ij}}{\partial x_j} + F_i &= \rho \frac{dv_i}{dt}, \\ \tau_{ij,j} + F_i &= \rho \dot{v}_i. \end{aligned} \quad (4.63)$$

In the second formulation, we use commas for derivatives with respect to spatial coordinates, just like in (4.2) and (4.3), and an over-dot for the total time derivative (Appendix 1).

The total derivative with respect to time of the velocity $\mathbf{v}$ can be expressed in the form,

$$\frac{dv_i}{dt} = \frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j}.$$

When using infinitesimal deformations and for very small velocities and changes in velocity with distance, second-order terms can be neglected and the total derivative with respect to time approximated by the partial derivatives. An over-dot on a letter will denote now a partial derivative with respect to time. If we substitute for the total the partial time

derivative in (4.62) and express the velocity in terms of displacement, replacing $\mathbf{T}$ by $\boldsymbol{\tau}$ (4.8), we obtain,

$$\int_V F_i dV + \int_S \tau_{ij} v_j dS = \int_V \rho \ddot{u}_i dV. \quad (4.64)$$

For an infinite medium, replacing the surface integral by one over the volume by means of Gauss's theorem, as in (4.62), we obtain

$$\frac{\partial \tau_{ij}}{\partial x_j} + F_i = \rho \frac{\partial^2 u_i}{\partial t^2}. \quad (4.65)$$

This is a very important formulation of the equation of motion that relates the accelerations with the forces and the spatial derivatives of the stresses. Equation (4.65) can be written in terms of the strain for an elastic medium by substitution of (4.16) into it:

$$\frac{\partial}{\partial x_j} (C_{ijkl} e_{kl}) + F_i = \rho \frac{\partial^2 u_i}{\partial t^2}. \quad (4.66)$$

In terms of the derivatives of displacements, according to (4.17), for constant elastic coefficients, we obtain

$$C_{ijkl} \frac{\partial^2 u_k}{\partial x_l \partial x_j} + F_i = \rho \frac{\partial^2 u_i}{\partial t^2}. \quad (4.67)$$

This is, then, a second-order differential equation for partial derivatives of the displacements with respect to the spatial coordinates and time, with an independent term formed by the body forces. If we specify these forces, solution of (4.67) gives us the elastic displacement field for the infinite medium. For an isotropic material, using equation (4.20) and substituting into (4.65) and using commas to indicate the derivatives, we obtain

$$\frac{\partial}{\partial x_j} [\lambda \delta_{ij} u_{k,k} + \mu (u_{i,j} + u_{j,i})] + F_i = \rho \ddot{u}_i. \quad (4.68)$$

For a homogeneous material, that is, for λ and μ constant, we can write the equation of motion in index and vector notation as

$$(\lambda + \mu) u_{k,ki} + \mu u_{i,jj} + F_i = \rho \ddot{u}_i,$$

$$(\lambda + \mu) \nabla(\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} + \mathbf{F} = \rho \ddot{\mathbf{u}}. \quad (4.69)$$

This expression represents the equation of motion in terms of displacements for a continuous, homogeneous, isotropic, infinite, elastic medium. The first term is the gradient of the divergence and the second the Laplacian of the displacements (Appendix 1). This equation is very important in seismology, since many problems can be solved using this approximation.

We can suppose that an earthquake is generated by processes that can be represented by a system of body forces acting in a certain focal region; then, outside this region, the only body forces are those of gravity ($\mathbf{F} = \mathbf{g}\rho$). However, except for waves with very large periods

($T > 600$ s), the influence of gravity is very small and is usually neglected (Section 5.5). Reducing the focal region to a point, we can use equation (4.69), approximating the Earth by an infinite medium, for the elastic displacements outside the focal region (Chapter 17).

The equation of motion (4.69) can also be expressed in terms of the cubic dilation θ and the rotation vector ω , whose relations to the displacements are given by (4.13) and (4.5), respectively. For this we use the equation that relates the curl of the curl of a vector to the gradient of the divergence and the Laplacian (Appendix 1). On substituting for the Laplacian of $\mathbf{u}$ in (4.69) its value in (A1.36), we obtain

$$(\lambda + 2\mu)u_{k,ki} - \mu e_{ijk} e_{klm} u_{n,lj} + F_i = \rho \ddot{u}_i. \quad (4.70)$$

Replacing θ and ω according to (4.13) and (4.5) and dividing by ρ results in

$$\begin{aligned} \alpha^2 \theta_{,i} - \beta^2 e_{ijk} \omega_{k,j} + \frac{F_i}{\rho} &= \ddot{u}_i, \\ \alpha^2 \nabla \theta - \beta^2 \nabla \times \omega + \frac{\mathbf{F}}{\rho} &= \ddot{\mathbf{u}}, \end{aligned} \quad (4.71)$$

or

$$\alpha^2 \nabla(\nabla \cdot \mathbf{u}) - \beta^2 \nabla \times (\nabla \times \mathbf{u}) + \mathbf{F}/\rho = \ddot{\mathbf{u}}.$$

In this equation we have introduced the parameters α and β whose values in terms of the elastic coefficients are

$$\alpha^2 = \frac{\lambda + 2\mu}{\rho} = \frac{K + \frac{3}{4}\mu}{\rho}, \quad (4.72)$$

$$\beta^2 = \mu/\rho. \quad (4.73)$$

The parameter α is related to θ and, in consequence, to changes in volume without changes in form, and β is related to ω , that is, to changes in form without changes in volume.

If, in equations (4.69) and (4.71), we make the forces $\mathbf{F}$ null, we obtain the homogeneous equation of motion, also known as *Navier's equation*:

$$(\lambda + \mu) \nabla(\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} = \rho \ddot{\mathbf{u}}, \quad (4.74)$$

$$\alpha^2 \nabla \theta - \beta^2 \nabla \times \omega = \ddot{\mathbf{u}}. \quad (4.75)$$

Solutions for these equations are the elastic displacements, and their cubic dilation and rotation, in a medium where there are no forces acting. With these equations we can study elastic perturbations in an infinite medium without considering the effects of any forces. We will see that these equations can be easily transformed into the wave equation (Chapter 5).

4.6 The Lagrangian formulation

Earthquakes are generated by the release of energy of tectonic origin by faulting that is propagated by seismic waves. In mechanics, the formulation of a problem from the point of view of energy is different from that of the equation of motion, which we have considered before. The study of the dynamics of a system from energy consideration can be done using the Lagrangian analytical dynamic formulation, proposed by Lagrange in 1788. Although this is not very often used in seismology it is important to recall here some of its basic principles in a simplified form.

The equation of energy continuity in a continuous medium in differential form, according to (4.60), is given by

$$\frac{\partial w}{\partial t} + \frac{\partial}{\partial x_i} (U_i w) = 0, \quad (4.76)$$

where w is the energy density and $\mathbf{U}$ is the velocity of the energy flux or propagation. The energy density in an elastic medium is given by the sum of kinetic T and potential V energies per unit mass,

$$W = T + V. \quad (4.77)$$

In terms of energy, the problem can be solved using the *Lagrangian density function*,

$$L = T - V. \quad (4.78)$$

For a discrete system, formed by N particles, this function depends on the generalized coordinates of each particle q_i , their time derivatives $\dot{q}_i$, the independent position variable x_i , and time, $L(q_i, \dot{q}_i, x_i, t)$. The application of *Hamilton's principle* of stationary action (named after William Rowan Hamilton) results in the *Lagrange equation* (Lindsay, 1960; Lanczos, 1986),

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0. \quad (4.79)$$

In a continuous system, which is the case for the problem of motion in an elastic medium, the Lagrangian density function depends on the *generalized coordinates* of each point of the continuous medium $q_i(x_k, t)$, their derivatives with respect to the spatial coordinates and time, and the independent variables x_i and t :

$$L\left(q_i, \frac{\partial q_i}{\partial t}, \frac{\partial q_i}{\partial x_k}, x_k, t\right).$$

In this case, application of the principle of stationary action leads to *Euler's equations* (summation is from 1 to 3) (Lindsay, 1960):

$$\sum_k \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial \left(\frac{\partial q_i}{\partial x_k} \right)} \right) + \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \left(\frac{\partial q_i}{\partial t} \right)} \right) - \frac{\partial L}{\partial q_i} = 0. \quad (4.80)$$

Equations (4.79) and (4.80) allow the solution of mechanical problems from the point of view of energy as an alternative to the application of Newton's second law.

The energy density w can be also expressed in terms of the Lagrangian function L , the generalized coordinates q_i , and their derivatives:

$$w = \sum_i \frac{\partial q_i}{\partial t} \frac{\partial L}{\partial q_i} - L. \quad (4.81)$$

By taking the time derivative of this expression, we obtain

$$\frac{\partial w}{\partial t} = \sum_i \left(\frac{\partial^2 q_i}{\partial t^2} \frac{\partial L}{\partial q_i} + \frac{\partial q_i}{\partial t} \frac{\partial}{\partial t} \frac{\partial L}{\partial q_i} \right) - \frac{\partial L}{\partial t}. \quad (4.82)$$

The flux of energy per unit surface and unit time H' (Section 4.5.1) is given by

$$H' = wU, \quad (4.83)$$

where U is the velocity of propagation or transport of energy. The flux of energy can also be derived from the Lagrangian function L (Lindsay, 1960):

$$H'_k = \sum_i \frac{\partial q_i}{\partial t} \frac{\partial L}{\partial \left(\frac{\partial q_i}{\partial x_k} \right)}. \quad (4.84)$$

An example of the application of these principles to the propagation of energy in shear motion in an elastic medium will be seen in the next chapter (Section 5.4).

4.7 Potential functions of displacements and forces

Displacements $\mathbf{u}(\mathbf{x}, t)$ in an elastic medium form a vector field. We can, therefore, apply *Helmholtz's theorem*, which allows their representation in terms of two potential functions, a scalar potential ϕ and a vector potential $\boldsymbol{\psi}$:

$$\begin{aligned} \mathbf{u} &= \nabla \phi + \nabla \times \boldsymbol{\psi} \\ u_i &= \phi_{,i} + e_{ijk} \psi_{k,j}. \end{aligned} \quad (4.85)$$

The vector potential $\boldsymbol{\psi}$ must satisfy the condition that its divergence is zero ($\nabla \cdot \boldsymbol{\psi} = 0$). Using expressions (4.13) and (4.5) it is easy to deduce the relations of the two potentials to the cubic dilation θ and the rotation $\boldsymbol{\omega}$:

$$\theta = \nabla^2 \phi, \quad (4.86)$$

$$\boldsymbol{\omega} = -\nabla^2 \boldsymbol{\psi}. \quad (4.87)$$

These relations indicate that ϕ is related to changes in volume and $\boldsymbol{\psi}$ to changes in form.

The body forces $\mathbf{F}$ can also be represented in a similar form by two potential functions, a scalar potential Φ and a vector potential of zero divergence Ψ :

$$\mathbf{F} = \nabla \Phi + \nabla \times \Psi. \quad (4.88)$$

Although the displacements and the forces may be defined for a point or a region, the potentials that represent them are defined for the entire space. Equation (4.71) can now be written in terms of the potentials, using (4.86)–(4.88), and can be separated into a scalar equation and a vector equation:

$$\alpha^2 \nabla^2 \phi + \frac{\Phi}{\rho} = \ddot{\phi} \quad (4.89)$$

$$\beta^2 \nabla^2 \psi + \frac{\Psi}{\rho} = \ddot{\psi}. \quad (4.90)$$

The vector differential equations of motion (4.69) and (4.71) are now expressed in terms of the potentials by a scalar and a vector equation of a much more simple form. We must remember that the three components of the vector potentials ψ and Ψ are not independent, since they must satisfy the condition that their divergences are zero. Then (4.89) and (4.90) represent only three independent equations corresponding to the three equations of (4.69) or (4.71). Equation (4.89) is related to the elastic perturbations which imply only changes in volume, and equation (4.90) is related to perturbations with changes in form only. This separation of the components of the equation of motion is very important, because it greatly facilitates its solution.

4.8 The Green and Somigliana functions of elastodynamics

In the Earth, neglecting the forces of gravity, body forces in the equation of motion (4.64) may be used to represent the processes that generate earthquakes. In general, these forces $\mathbf{F}(\mathbf{x}, t)$, functions of the spatial coordinates and time, may be different for each earthquake and are defined only inside a certain volume we have called the focal region. An example of a convenient simple time dependence is the harmonic function

$$\mathbf{F}(\mathbf{x}, t) = \mathbf{F}(\mathbf{x})e^{i\omega t}. \quad (4.91)$$

This form of time dependence simplifies solution of many problems in seismology. Use of the harmonic function is not in itself a very realistic way to represent the time dependence of forces that generate earthquakes, but solutions to this problem may be used to find solutions for other time functions by means of the Fourier transform (Appendix 4).

A type of body force of great importance in the solution of many problems of elastodynamics is that formed by a unit impulsive force acting at a point in space and time with an arbitrary direction. This force may be represented mathematically by means

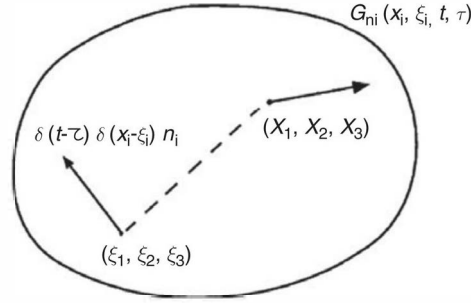


Fig. 4.7

Displacements G_{ni} (Green's function) at x_i produced by an impulsive force in time and space at ξ_i .

of Dirac's delta function $\delta(x)$ (used already in Section 3.2), which has the important property,

$$f(x_0) = \int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx. \quad (4.92)$$

The impulsive force can be represented by

$$F_i(x_s, t) = \delta(x_s - \xi_s) \delta(t - \tau) \delta_{in}. \quad (4.93)$$

The force is applied at the point of the coordinates ξ_i and the time τ , and is null outside this point and time. Its orientation is given by its three components represented by the index n . If we substitute this force into the equation of motion (4.64), its solutions are the elastic displacements for every point of coordinates x for every time t , in a certain volume V , surrounded by a surface S . On replacing the stress in terms of the derivatives of the displacements (4.17) into the equation of motion (4.64), we obtain

$$\int_V \rho \ddot{G}_{ni} dV - \int_S C_{ijkl} G_{nk,l} v_j dS = \int_V \delta(x_s - \xi_s) \delta(t - \tau) \delta_{in} dV. \quad (4.94)$$

Each component of the displacement (index i) depends on the orientation of the force (index n) and thus the displacement is given by a second-order tensor with components $G_{ni}(x_s, \xi_s, t, \tau)$, which is a function of the coordinates and time (x_s, t) of each point in V and of the coordinates and time of the point of application of the force (ξ_s, τ) (Fig. 4.7).

On taking the reciprocal of the surface integral by using Gauss's theorem, as in (4.62), for a homogeneous, infinite medium according to (4.67), we have the differential equation

$$\rho \ddot{G}_{ni} - C_{ijkl} G_{nk,lj} = \delta(x_s - \xi_s) \delta(t - \tau) \delta_{in}. \quad (4.95)$$

The solutions of equations (4.94) and (4.95) represent the elastic displacements due to a unit impulse force acting at a point in space and time. For this reason the tensor $\mathbf{G}$ is called, after George Green, the *Green tensor* or *function* of elastodynamics, that is, the response of the medium to an impulsive excitation. The form of this function depends on the characteristics of the medium, its elastic coefficients, and its density. In a finite medium (4.94), it depends also on the shape and size of the volume V and the boundary conditions on its surface S . For

each medium there is a different Green function that defines how this medium reacts mechanically to an impulsive excitation force (the solution for an infinite medium is given in [Chapter 17](#)).

For the static case, that is, when there is no time dependence, equation (4.94) reduces to

$$\int_S C_{ijkl} S_{ni,l} dS = - \int_V \delta(x_s - \xi_s) \delta_{ni} dV, \quad (4.96)$$

and for an infinite homogeneous medium according to (4.95),

$$C_{ijkl} S_{nk,lj} = -\delta(x_s - \xi_s) \delta_{ni}, \quad (4.97)$$

where the tensor S_{ni} is called, after Carlo Somigliana, the *Somigliana tensor* or *function*, which represents the static displacements produced by an impulsive unit force acting at a point of place and time. The Green and Somigliana tensors or functions represent, therefore, the mechanical response to an impulsive excitation and are proper characteristics of each medium.

4.9 Theorems of reciprocity and representation

As we have seen, displacements in an elastic medium depend on the body forces and stresses acting on it. Let us consider an elastic medium of volume V surrounded by a surface S . For a system of body forces $\mathbf{f}$ acting on each volume element dV and stresses $\mathbf{T}^u$ on every element of surface dS , the displacements are $\mathbf{u}$. In the same volume let us consider also a different system of forces and stresses $\mathbf{g}$ and $\mathbf{T}^w$, to which correspond the displacements $\mathbf{w}$ ([Fig. 4.8](#)). For each case we can write the equation of motion (4.64) in the form,

$$\int_V (f_i - \rho \ddot{u}_i) dV + \int_S T_i^u dS = 0. \quad (4.98)$$

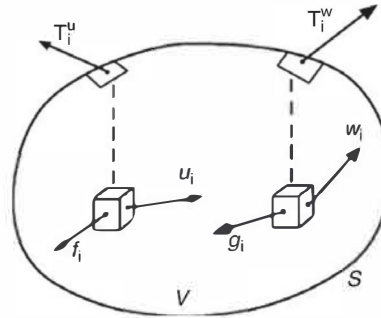


Fig. 4.8

Elastic displacements $\mathbf{u}$ and $\mathbf{w}$ corresponding to body forces $\mathbf{f}$ and $\mathbf{g}$ acting inside the volume V and stresses $\mathbf{T}^u$ and $\mathbf{T}^w$ acting on the surface S .

A similar expression can be written for $\mathbf{g}$, $\mathbf{T}^w$, and $\mathbf{w}$. In (4.98) we take the scalar product of each term with $\mathbf{w}$ and, in its analog, we do this with $\mathbf{u}$. Since each expression is now a scalar equation equal to zero, we can equate them, resulting in

$$\int_V (f_i - \rho \ddot{u}_i) w_i dV + \int_S T_i^u w_i dS = \int_V (g_i - \rho \ddot{w}_i) u_i + \int_S T_i^w u_i dS. \quad (4.99)$$

This expression is known, after Enrico Betti, as *Betti's reciprocity theorem*. It shows the reciprocal relation between the displacements corresponding to two systems of forces and stresses acting in the same volume. If we reorder the terms in (4.99) and integrate each one through all time, we obtain

$$\int_{-\infty}^{\infty} dt \int_V \rho (u_i \dot{w}_i - w_i \dot{u}_i) dV = \int_{-\infty}^{\infty} dt \int_V (u_i g_i - w_i f_i) dV + \int_{-\infty}^{\infty} dt \int_S (u_i T_i^w - w_i T_i^u) dS. \quad (4.100)$$

In this expression, the products are convolutions in time (Appendix 4). Thus, the first term can be written in full form by changing the order of integration:

$$\int_V dV \int_{-\infty}^{\infty} \rho [u_i(t) \dot{w}_i(\tau - t) - \dot{u}_i(t) w_i(\tau - t)] dt. \quad (4.101)$$

Equation (4.100) is known as the *Green–Volterra formula* (after Green and Vito Volterra) and is a generalization of Betti's reciprocity theorem. This expression relates the displacements and accelerations produced by two systems of forces and stresses on the same volume integrated over space and time.

A particular case results if displacements and velocities are null for all times previous to a given time. This condition implies the *causality principle*, namely that the medium is at rest until a given time when motion starts, which is the case in the occurrence of earthquakes. If this time is $t = 0$, it follows that

$$u_i = \dot{u}_i = 0, t \leq 0$$

$$w_i = \dot{w}_i = 0, t \leq 0.$$

Under these conditions, the integral of time in (4.101) can be written in the form,

$$\rho \int_0^{\tau} \frac{d}{dt} [\dot{u}_i(t) w_i(\tau - t) + \dot{w}_i(\tau - t) u_i(t)] dt. \quad (4.102)$$

The change in sign is due to the fact that the first derivative with respect to time of $\mathbf{w}$ is negative. On integrating we obtain

$$\rho [\dot{u}_i(\tau) w_i(0) + \dot{w}_i(0) u_i(\tau) - \dot{u}_i(0) w_i(\tau) - \dot{w}_i(\tau) u_i(0)].$$

Taking into account the conditions specified above for $\mathbf{u}$, $\mathbf{w}$, and their derivatives at $t = 0$, this expression is zero (Aki and Richards, 1980, 28–31). Then, equation (4.100) becomes

$$\int_{-\infty}^{\infty} dt \int_V (u_i g_i - w_i f_i) dV = \int_{-\infty}^{\infty} dt \int_S (w_i T_i^u - u_i T_i^w) dS. \quad (4.103)$$

This is a very important result for seismology, since it allows the representation of the displacements due to a system of forces by those produced by a different system, given that

causality conditions are satisfied. We can, then, represent the displacements due to a complicated system of forces in terms of those produced by a simpler one.

We can select, as the simplest system of forces, an impulsive force in space and time. As we have seen above, the displacements corresponding to this type of force are given by Green's function. Thus, we substitute into (4.103) $\mathbf{g}$ by expression (4.93) and the displacements $\mathbf{w}$ by Green's tensor $\mathbf{G}$. Accordingly, equation (4.103) becomes

$$\begin{aligned} \int_{-\infty}^{\infty} dt \int_V [u_i \delta(x_s - \xi_s) \delta(t - \tau) \delta_{in} - G_{ni} f_i] dV \\ = \int_{-\infty}^{\infty} dt \int_S (G_{ni} T_i - u_i C_{ijkl} G_{nk,l} v_j) dS, \end{aligned} \quad (4.104)$$

where we have replaced $\mathbf{T}^u$ by $\mathbf{T}$ and $\mathbf{T}^w$ by its value in terms of the derivatives of the displacements (4.17),

$$T_i^w = \tau_{ij} v_j = C_{ijkl} G_{nk,l} v_j. \quad (4.105)$$

According to the property of the delta function (4.92), the first integral of (4.104) results in

$$\int_{-\infty}^{\infty} dt \int_V u_i(x_s, t) \delta(x_s - \xi_s) \delta(t - \tau) \delta_{in} dV = u_n(\xi_s, \tau). \quad (4.106)$$

We replace (4.106) in (4.104) and since equation (4.104) is symmetric with respect to the variables $\mathbf{x}$, t , ξ , and τ , we can interchange them. Thus we obtain

$$u_n(x_s, t) = \int_{-\infty}^{\infty} d\tau \int_V f_i G_{ni} dV + \int_{-\infty}^{\infty} d\tau \int_S [G_{ni} T_i - C_{ijkl} u_i G_{nk,l} v_j] dS. \quad (4.107)$$

This is an important result, as a special case of (4.103), and is generally referred to as the *representation theorem*. This equation gives us the elastic displacements inside a volume V as the sum of two double integrals over time and space. The first is the volume integral of the body forces multiplied by Green's function. The second is the surface integral of the stress multiplied by Green's function minus the displacements times the derivatives of Green's function. The displacements $\mathbf{u}(\mathbf{x}, t)$ are given for every point of the volume V at every time, provided that the causality principle is satisfied. In the two integrals $\mathbf{f}$, $\mathbf{T}$, and $\mathbf{u}$ are functions of the integration variables ξ_i and τ , and, in general, are also $v_i(\xi_s)$, $dV(\xi_s)$, and $dS(\xi_s)$. Green's function $G_{ni}(\xi_s, \tau; x_s, t)$ depends both on x_s and t , and on ξ_s and τ (Fig. 4.9).

Equation (4.107) allows us to determine the elastic displacements produced by a system of body forces defined in a volume and/or by a system of stresses and displacements defined over a surface by means of the Green function, as we will see in Chapter 17. To use this equation we need to have previously determined Green's function for the medium in question with specified characteristics (dimensions, shape, density, and elastic coefficients). This means that we need only solve the equation of motion once to find Green's function. Once Green's function for a certain medium is known, we can determine the elastic displacements for any other type of system of body forces in a volume and/or of stresses and displacements on a surface using equation (4.107). In this equation, Green's function acts as a *propagator*, since it propagates the effects of forces, stresses, or displacements

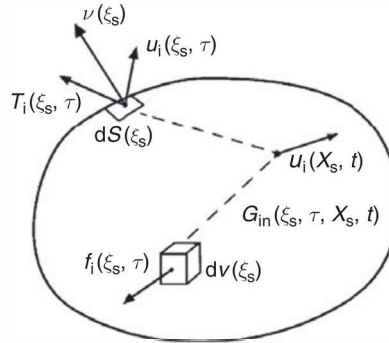


Fig. 4.9

Elastic displacements $\mathbf{u}$ at point $\mathbf{x}$ and time t , corresponding to a body force $\mathbf{f}$ and stress $\mathbf{T}$ acting at ξ and τ , in a medium with Green's function $\mathbf{G}$.

defined on coordinates (ξ, τ) to determine the elastic displacements at points of the coordinates $(\mathbf{x}, t)$.

In general, determination of the Green function is not easy. The difficulty increases with the complexity of the medium. The simplest case is for an infinite, homogeneous, isotropic, perfectly elastic medium (Chapter 17). Solutions for other media such as a layered half-space and a radially symmetric sphere have been determined and are often used. The advantage of using equation (4.107) is that, for a specified medium, Green's function, which implies the solution of the equation of motion, need only be solved once. As we will see in Chapter 17, equation (4.107) is very important in the determination of the source mechanism of earthquakes, allowing convenient use of different types of source models once the Green function for the medium is known.

4.10 Summary

To formalize and quantify the study of earthquakes we apply the principles and equations of the mechanics of an elastic medium to the motions produced in the Earth. The basic concepts are the ground displacements, velocity, and acceleration, its deformation or strain, and the applied forces and stresses. For a linear elastic body the relation between stress and strain is given by Hooke's law through the elastic coefficients. Stresses and forces that produce deformations result in work done in the medium with heat production so that energy and heat considerations are to be considered.

The basic equations which describe the behavior of an elastic medium are the continuity of mass and energy and the equation of motion or momentum, that is, the application of Newton's second law. This last equation (for example in differential form (4.65)) gives the fundamental relation between the acting forces and stresses and the produced displacements in a medium. An alternative formulation in terms of the energy is given by the Lagrangian. For solution of the equation of motion we have introduced the potential of displacements and forces, the Green and Somigliana functions, and the theorems of

reciprocity and representation. These basic concepts and equations of mechanics give us the necessary tools to study quantitatively the motion produced in the Earth by earthquakes.

4.11 Problems

- P4.1 The coordinate system x'_1, x'_2, x'_3 is obtained by rotation of the system x_1, x_2, x_3 by 90° around x_1 so that $x'_2 = x_3$ and $x'_3 = -x_2$. If the system x_1, x_2, x_3 corresponds to the principal strain axes, show that: $e'_{23} = -e_{32}$, $e'_{22} = e_{33}$, and $e'_{33} = e_{22}$.

- P4.2 Given the stress tensor

$$\begin{pmatrix} 2 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

find the principal stresses, the principal axes, the invariants I_1, I_2 , and I_3 , the deviatoric stress tensor, and the invariants J_2 and J_3 .

- P4.3 Show that the invariants of the deviatoric stress tensor in terms of its principal values are

$$J_1 = 0; \quad J_2 = \frac{1}{2}(s_1^2 + s_2^2 + s_3^2); \quad J_3 = \frac{1}{2}(s_1^3 + s_2^3 + s_3^3).$$

- P4.4 Show that the relation between the invariants of the stress tensor and those of its deviatoric tensor are

$$J_2 = 3\sigma_0 + I_2; \quad J_3 = I_3 - \sigma_0 I_2 + 2\sigma_0^3.$$

- P4.5 Given a stress tensor

$$\begin{pmatrix} 3x_1x_2 & 5x_2^2 & 0 \\ 5x_2^2 & 0 & 2x_3 \\ 0 & 2x_3 & 0 \end{pmatrix},$$

determine the stress vector T at point $(2, 1, \sqrt{3})$ through a plane tangential to the cylindrical surface $x_2^2 + x_3^2 = 4$.

- P4.6 Given the stress tensor

$$\begin{pmatrix} 1/4 & -5/4 & (1/2)^{3/2} \\ -5/4 & 1/4 & -(1/2)^{3/2} \\ (1/2)^{3/2} & -(1/2)^{3/2} & 3/2 \end{pmatrix},$$

- determine the principal stresses (eigenvalues);
- determine the angles with the coordinate axes of the eigenvector corresponding to the greatest principal stress.

- P4.7 Show that e_{ij} and ω_{ij} are symmetric and antisymmetric tensors and that $e_{ij}\omega_{ij} = 0$.

P4.8 Derive from the stress–strain relation for an isotropic medium that

$$\lambda = \frac{\sigma E}{(1 + \sigma)(1 - 2\sigma)}; \quad \mu = \frac{E}{2(1 + \sigma)},$$

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}; \quad \sigma = \frac{\lambda}{2(\lambda + \mu)},$$

where σ is Poisson's ratio.

P4.9 Derive for the stress–strain relation for an isotropic medium that

$$e_{11} = \frac{(\lambda + \mu)\tau_{11}}{\mu(3\lambda + 2\mu)} - \frac{\lambda(\tau_{22} + \tau_{33})}{2\mu(3\lambda + 2\mu)}.$$

As mentioned in [Chapter 2](#), earthquakes produce waves that propagate through the Earth, cause damage at short distances, and are recorded and measured by seismographs, as seen in [Chapter 3](#). To understand the phenomenon of wave propagation in the Earth we begin with the simplest case of waves propagating in an infinite elastic medium. As we saw in [Chapter 4](#), in an infinite, homogeneous, isotropic, elastic medium, the basic equations of mechanics are the equation of motion and of continuity and propagation of energy. Under certain conditions these equations lead to the phenomenon of wave propagation that can be formulated in terms of *wave equations*, first proposed in 1747 by Jean d'Alembert. In this chapter we present the solutions for the wave equation in terms of plane, spherical, and cylindrical waves in an infinite elastic medium, their characteristics, and relations with the observations.

5.1 Wave equations for an elastic medium

In order to study the propagation of waves in an infinite elastic medium, we begin with Navier's equation which can be expressed in terms of the displacements (4.74) and of the cubic dilation and rotation vector (4.75) equations, which may be easily transformed into wave equations. We begin with application of the divergence operation in equation (4.75). The divergence of the gradient of θ is its Laplacian, that of the curl of ω is null, and the divergence of the displacement $\mathbf{u}$ is the cubic dilation θ . Thus, we obtain

$$\nabla^2 \theta = \frac{1}{\alpha^2} \frac{\partial^2 \theta}{\partial t^2}. \quad (5.1)$$

To the same equation (4.75) we apply the curl operation. The curl of the gradient of the scalar function θ is null and that of the displacement $\mathbf{u}$ is the rotation vector ω . The curl of the curl of ω is equal to the gradient of the divergence, which is null minus the Laplacian (A1.30). The result is

$$\nabla^2 \omega = \frac{1}{\beta^2} \frac{\partial^2 \omega}{\partial t^2}. \quad (5.2)$$

The same results (5.1) and (5.2) can be obtained by taking the divergence and the curl operations in equation (4.74). Equations (5.1) and (5.2) have the form of wave equations for the scalar function θ and vector function ω . The solutions of both equations represent waves that propagate in the elastic medium and the parameters α and β are their velocities. These velocities are functions of the elastic coefficients λ and μ and the density ρ , as we have

already found in (4.72) and (4.73). Velocity α is always greater than β , for example, if Poisson's ratio (4.29) is 0.25, $\alpha = \sqrt{3}\beta$. Because θ represents changes in volume without changes in shape, solutions of equation (5.1) correspond to compressional and dilational motion, or *longitudinal waves*. As waves propagate, the elastic material expands and contracts keeping the same form. In seismology, these waves are given the name of P (*prima*) waves, since they are the first to be observed in the seismograms (α is greater than β). Solutions of equation (5.2) represent *shear waves* that propagate with velocity β . In these waves the medium changes in shape, but not in volume, since the divergence of $\mathbf{u}$ is null. In seismology these waves are given the name of S (*secunda*) waves, because they are the second type of prominent waves arriving after P waves. Therefore, in an infinite, homogeneous, isotropic, elastic medium, there exist only these two types of waves that are called *body waves*, because they travel through the body of the medium. This important result was first found by Poisson in 1829. For a liquid medium, $\mu = 0$, and there are only P waves.

The same result can be obtained for the potentials ϕ and ψ defined in (4.85). If in equations (4.89) and (4.90) we make null the potentials of the forces Φ and Ψ , we obtain

$$\nabla^2 \phi = \frac{1}{\alpha^2} \frac{\partial^2 \phi}{\partial t^2}, \quad (5.3)$$

$$\nabla^2 \psi = \frac{1}{\beta^2} \frac{\partial^2 \psi}{\partial t^2}. \quad (5.4)$$

Therefore, under conditions of the absence of body forces, the potentials ϕ and ψ are also solutions of the wave equation. Since α and β are the velocities of P and S waves, ϕ is the potential of P waves and ψ that of S waves. The total elastic displacement $\mathbf{u}$ is the sum of the displacements of P and S waves and can be written as

$$\mathbf{u} = \mathbf{u}^P + \mathbf{u}^S. \quad (5.5)$$

According to equation (4.85),

$$\mathbf{u}^P = \nabla \phi, \quad (5.6)$$

$$\mathbf{u}^S = \nabla \times \psi. \quad (5.7)$$

Displacements of the P and S waves can be deduced from the potentials ϕ and ψ , respectively.

Elastic displacements $\mathbf{u}$ in the absence of body forces are solutions of equation (4.74), which is not a wave equation. However, if they represent only P or S waves, they are solutions of wave equations. For example, if $\mathbf{u}$ depends only on x_1 and has only a component in the direction of x_2 , that is, $u_2(x_1, t)$, corresponding to a transversal displacement, then, on substituting into equation (4.74), we obtain

$$\frac{\partial^2 u_2}{\partial x_1^2} = \frac{1}{\beta^2} \frac{\partial^2 u_2}{\partial t^2}. \quad (5.8)$$

The displacement u_2 is in this case the solution of the wave equation. Solutions of (5.8) correspond to S waves that propagate in the x_1 direction with velocity β .

The same result can be obtained for a displacement of only a component $u_1(x_1, t)$ that depends only on x_1 . By substitution into (4.74) we obtain

$$\frac{\partial^2 u_1}{\partial x_1^2} = \frac{1}{\alpha^2} \frac{\partial^2 u_1}{\partial t^2}. \quad (5.9)$$

Therefore u_1 is, in this case, the solution of the wave equation. The solutions are P waves with displacements and propagation in the same direction x_1 and velocity α .

If, in the wave equation (5.3), the potential ϕ has a harmonic dependence on time with angular frequency $\omega = 2\pi f$, where f is the frequency, $f = 1/T$, and T the period,

$$\phi(x, t) = \phi(x) e^{-i\omega t}.$$

After substitution, we obtain

$$(\nabla^2 + k_a^2)\phi = 0, \quad (5.10)$$

where $k_a = \omega/\alpha$ is the wave number of P waves. This is the *Helmholtz equation*, from which the time dependence has been eliminated. The same can be done starting from equation (5.4) and we obtain an equation for ψ similar to (5.10), in which $k_\beta = \omega/\beta$ is the wave number of S waves.

5.2 Solutions of the wave equation

From the equation of motion for displacements in an elastic medium in the absence of body forces (4.74), we have derived wave equations for the cubic dilation θ (5.1), the rotation vector ω (5.2), and the scalar ϕ (5.3) and vector ψ (5.4) potentials. The simplest case of a wave equation is that in one dimension:

$$\frac{\partial^2}{\partial x^2} f(x, t) = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} f(x, t), \quad (5.11)$$

where c is the velocity of wave propagation. To solve this equation we apply the method of separation of variables, making $f(x, t) = R(x)T(t)$. By substitution into (5.11), we obtain

$$\frac{c^2}{R} \frac{d^2 R}{dx^2} = \frac{1}{T} \frac{d^2 T}{dt^2} = -\omega^2, \quad (5.12)$$

where we have introduced ω^2 as the variables separation constant. From (5.12) we obtain two separate equations for R and T :

$$\frac{d^2 R}{dx^2} + k^2 R = 0, \quad (5.13)$$

$$\frac{d^2 T}{dt^2} + \omega^2 T = 0, \quad (5.14)$$

where $k^2 = \omega^2/c^2$. The solutions of both equations can be expressed in harmonic functions of x and t and, for $f(x, t)$, their product is

$$f(x, t) = Ae^{i(kx - \omega t)} + Be^{i(kx + \omega t)}. \quad (5.15)$$

From the form of this solution, we can see that the separation constant ω represents the *angular frequency* and k is the *wave number*. The velocity of wave propagation is $c = \omega/k$. In conclusion, equation (5.15) represents harmonic waves of frequency ω that propagate with velocity c in the positive and negative x directions.

A more general solution of the wave equation, not restricted to harmonic functions, can be written in the form,

$$f(x, t) = f(x - ct) + f(x + ct), \quad (5.16)$$

where, as before, $c = \omega/k$ is the velocity of propagation and, depending on the sign, waves propagate in the positive or negative x direction.

A particular solution, such as that in (5.15), may be given by harmonic functions, represented by imaginary exponentials, sines, and cosines. For one value of ω , they represent monochromatic waves, which, for propagation in the positive x direction, may be expressed in the following forms:

$$f(x, t) = C e^{i(kx - \omega t + \varepsilon)}, \quad (5.17)$$

$$f(x, t) = C \cos[k(x - ct) + \varepsilon], \quad (5.18)$$

$$f(x, t) = A \cos(kx - \omega t) + B \sin(kx - \omega t). \quad (5.19)$$

In (5.17) and (5.18), C is the *amplitude* of the wave and ε is the *phase at the origin* or the *initial phase*. In (5.19), A and B include the amplitude and initial phase. In all cases, the solutions have two arbitrary constants, A and B or C and ε . Their relation is

$$C^2 = A^2 + B^2, \quad (5.20)$$

$$\varepsilon = \tan^{-1}(B/A). \quad (5.21)$$

For harmonic waves, $f(x, t)$ has the same values at constant intervals of distance, multiples of the *wave length* $\lambda = 2\pi/k$, and at time multiples of the *period* $T = 2\pi/\omega$ (Fig. 5.1). In terms of λ and T , equation (5.18) may be written in the form,

$$f(x, t) = C \cos \left[2\pi \left(\frac{x}{\lambda} - \frac{t}{T} \right) + \varepsilon \right].$$

The argument ζ of the harmonic function for each value of x and t is the *phase*

$$\zeta = k(x - ct) + \varepsilon = kx - \omega t + \varepsilon = 2\pi \left(\frac{x}{\lambda} - \frac{t}{T} \right) + \varepsilon. \quad (5.22)$$

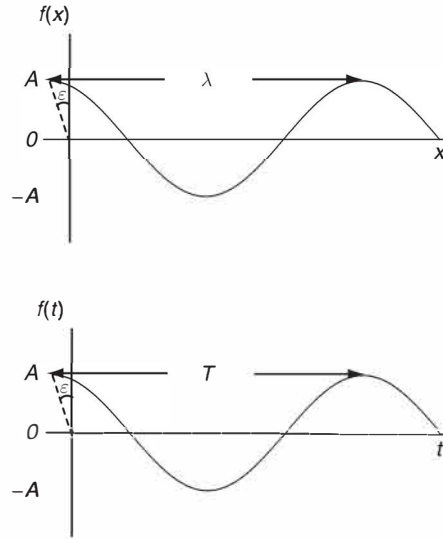


Fig. 5.1 A representation of sinusoidal wave motion as a function of the distance x and time t . A is the amplitude, ε is the initial phase, λ is the wave length, and T is the period.

For a constant value of ζ , if we vary t , x also varies, and this phase value propagates in time and space with a velocity c . For this reason, c is the velocity of propagation of each phase or *phase velocity*.

5.2.1 Wave fronts and rays

In three dimensions, a solution of the wave equation for monochromatic waves of frequency ω may be written as

$$f(x_j, t) = A \exp \{i[kS(x_j) - \omega t + \varepsilon]\}. \quad (5.23)$$

The geometric surface of all points in space for which the phase ζ has the same constant value forms a *wave front*. According to (5.23), for each time, the function $S(x_j)$ gives us the spatial coordinates of the corresponding wave front. Let us suppose that the phase is $\zeta = 0$, then the spatial coordinates of the wave front corresponding to this value of the phase for times t_1 and t_2 are given by (Fig. 5.2)

$$S_1(x_j) = \frac{\omega}{k} t_1 - \frac{\varepsilon}{k}, \quad (5.24)$$

$$S_2(x_j) = \frac{\omega}{k} t_2 - \frac{\varepsilon}{k}. \quad (5.25)$$

The unit vector normal at each point of the wave front at each time defines the *ray's direction* or direction of propagation of each element of the wave front. The trajectory or path followed by the normal of the same element of the wave front during a certain time

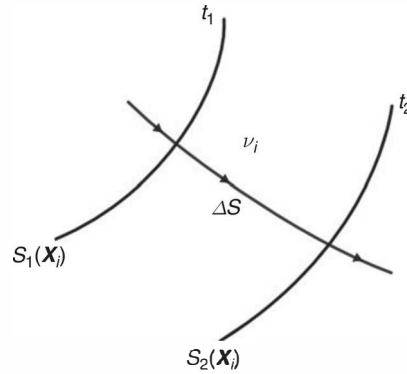


Fig. 5.2 Wave fronts $S_1(x)$ and $S_2(x)$ corresponding to times t_1 and t_2 and the ray's trajectory. ΔS is an element of distance along the ray.

interval defines the *ray's trajectory* or *path*. The orientation of the ray at each time is given by its direction cosines which can be derived from the gradient of S :

$$v_i = \frac{\frac{\partial S}{\partial x_i}}{\sqrt{\left(\frac{\partial S}{\partial x_1}\right)^2 + \left(\frac{\partial S}{\partial x_2}\right)^2 + \left(\frac{\partial S}{\partial x_3}\right)^2}}. \quad (5.26)$$

The equations of ray trajectories are given by

$$\frac{dx_1}{\frac{\partial S}{\partial x_1}} = \frac{dx_2}{\frac{\partial S}{\partial x_2}} = \frac{dx_3}{\frac{\partial S}{\partial x_3}}, \quad (5.27)$$

where $ds^2 = dx_1^2 + dx_2^2 + dx_3^2$ is the element of distance in the ray's direction. If the wave fronts for the same phase, which correspond to times t and $t + \Delta t$, are, respectively, S_1 and S_2 , then, from (5.24) and (5.25), we obtain

$$kS_1 - \omega t = kS_2 - \omega(t + \Delta t). \quad (5.28)$$

If the increment in the direction of propagation between the two wave fronts is $S_2 - S_1 = \Delta S$, then, from (5.28), we obtain

$$\frac{\Delta S}{\Delta t} = \frac{\omega}{k} = c, \quad (5.29)$$

where c is, again, the velocity of propagation of the wave front corresponding to the same phase or *phase velocity*. According to (5.29), the wave front advances with this velocity a distance ΔS in a time Δt . This velocity may depend on the spatial coordinates and in this case each element of the wave front advances with a different velocity in a different direction and its form will change.

5.2.2 Waves of several frequencies

Up to this point, we have considered only monochromatic harmonic waves, that is, harmonic waves with only one value of the frequency ω . This is a simplification since earthquake waves are not monochromatic. Waves with an arbitrary dependence on time can be represented by the sum or integral of harmonic waves of different frequencies using Fourier's theorem ([Appendix 4](#)):

$$f(x_i, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp \left[i \left(\frac{\omega}{c(\omega)} S(x_i) - \omega t \right) \right] d\omega, \quad (5.30)$$

where $F(\omega)$ is a complex function that is called the *complex spectrum* of $f(x, t)$ and can be represented as

$$F(\omega) = R(\omega) + iI(\omega) = A(\omega)e^{i\Phi(\omega)}, \quad (5.31)$$

where

$$A(\omega) = \left[R(\omega)^2 + I(\omega)^2 \right]^{1/2},$$

$$\Phi(\omega) = \tan^{-1}[I(\omega) + R(\omega)].$$

$A(\omega)$ is the *amplitude spectrum* and $\Phi(\omega)$ is the *phase spectrum*. These two functions of frequency represent the contributions of amplitude and initial phase by each harmonic component to the resulting function of time. The phase velocity $c(\omega)$ is now a function of frequency, that is, each harmonic component may have a different phase velocity. Motion at a particular point due to a propagating wave is given by a function of time, $f(t)$. The same information can be represented as a function of frequency, $F(\omega)$ that is called the spectrum of $f(t)$, which is obtained by means of the inverse Fourier transform ([A4.12](#)):

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt. \quad (5.32)$$

$f(t)$ is a real function that represents the amplitudes and phases of waves at each time for a particular point of space. Its Fourier transform $F(\omega)$ is a complex function that can be represented by its amplitudes $A(\omega)$ and initial phases $\Phi(\omega)$ for each frequency ([5.31](#)). In this way, problems of wave propagation can be studied in the time domain $f(t)$ or in the frequency domain $F(\omega)$ related by equations ([5.30](#)) and ([5.32](#)). Spectra show the contribution of each frequency to the waves observed.

5.3 Displacement, velocity, and acceleration

In the one-dimensional case, the displacement in an elastic medium due to a monochromatic wave that propagates in the positive direction of the x axis with velocity c is given by

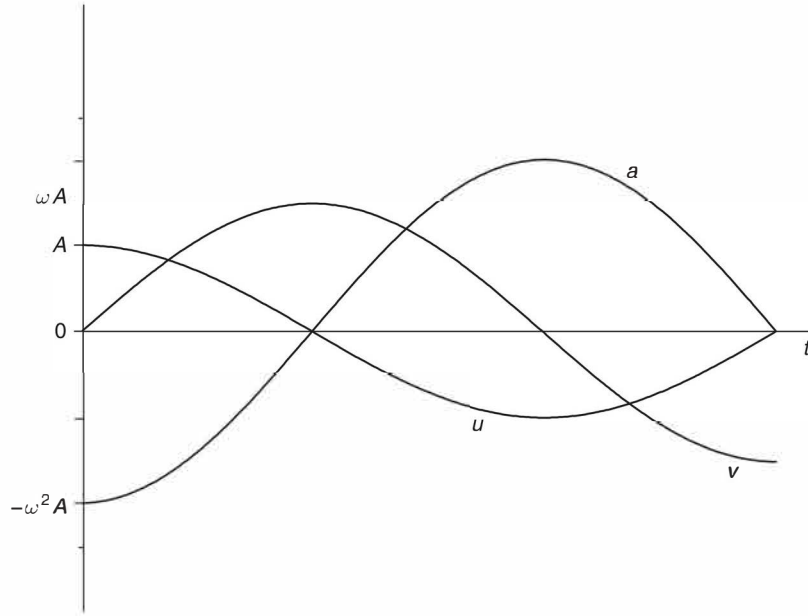


Fig. 5.3 The displacement, velocity, and acceleration for a cosine wave.

$$u(x, t) = A \cos \left[\omega \left(\frac{x}{c} - t \right) + \varepsilon \right]. \quad (5.33)$$

At a distance x , the displacement varies with time from A to $-A$ and the maxima correspond to phase values of $0, \pi, 2\pi, 3\pi$, etc. (Fig. 5.3). The velocity of this motion, or the *particle velocity* (not to be mistaken for the phase velocity), is given by the time derivative,

$$v(x, t) = \frac{\partial u}{\partial t} = A\omega \sin \left[\omega \left(\frac{x}{c} - t \right) + \varepsilon \right]. \quad (5.34)$$

The acceleration is given by the second derivative,

$$a(x, t) = \frac{\partial^2 u}{\partial t^2} = -A\omega^2 \cos \left[\omega \left(\frac{x}{c} - t \right) + \varepsilon \right]. \quad (5.35)$$

The particle velocity is shifted in phase by $\pi/2$ with respect to the displacement and the acceleration by π (Fig. 5.3). The amplitude of the velocity is multiplied by the frequency and that of the acceleration by its square. The study of ground displacement, velocity, and acceleration is an important subject in seismology. Damage produced by earthquakes is caused by the effects of such motion, especially the effect of its acceleration on buildings and other structures (Chapter 2). Recorded by seismographs, ground displacement, velocity, and acceleration constitute the basic observations for the quantitative study of earthquakes (Chapter 3).

5.4 The propagation of energy

Another approach that leads to the phenomenon of wave propagation is from the point of view of propagation of energy. This can be shown by a simple application of the principles of the Lagrangian formulation (Section 4.6) to the case of pure shear motion in an elastic medium. Let us consider a displacement u normal to the direction of propagation in the x direction. In this case in the equations of Section 4.6, $x_i = x$ and $\mathbf{u}_i = u$. The kinetic T and potential V energy densities are given by

$$T = \frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2, \quad (5.36)$$

$$V = \frac{\mu}{2} \left(\frac{\partial u}{\partial x} \right)^2. \quad (5.37)$$

The Lagrangian function (4.78) is

$$L = \frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2 - \frac{\mu}{2} \left(\frac{\partial u}{\partial x} \right)^2. \quad (5.38)$$

On applying Euler's equation (4.80), we obtain

$$\frac{\partial}{\partial x} \left(-\mu \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial t} \left(\rho \frac{\partial u}{\partial t} \right) = 0. \quad (5.39)$$

From this equation, we finally derive

$$\frac{\partial^2 u}{\partial x^2} = \frac{\rho}{\mu} \frac{\partial^2 u}{\partial t^2}. \quad (5.40)$$

This is the wave equation for a shear displacement or S wave (5.8), where, as before, $\beta^2 = \mu/\rho$ is the velocity of phase propagation. Here we have derived the wave equation from energy considerations by applying Euler's equation. Previously ((5.1) and (5.2)), we had obtained the same result from the equation of motion.

The energy density w , according to (4.77), is the sum of the kinetic and potential energies:

$$w = \frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2 + \frac{\mu}{2} \left(\frac{\partial u}{\partial x} \right)^2, \quad (5.41)$$

and its time derivative is

$$\frac{\partial w}{\partial t} = \rho \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} + \mu \frac{\partial u}{\partial x} \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial x} \right). \quad (5.42)$$

If we replace the second time derivative of u according to (5.40) and change the order of derivatives in (5.42), we obtain

$$\frac{\partial w}{\partial t} = \mu \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right). \quad (5.43)$$

Substituting this expression into the energy-continuity equation (4.76) results in

$$\frac{\partial}{\partial x} \left(\mu \frac{\partial u}{\partial t} \frac{\partial u}{\partial x} + wU \right) = 0. \quad (5.44)$$

The quantity inside the brackets is independent of x . For a localized wave whose value tends to zero when x tends to infinity, this quantity is zero for all times. Hence, the velocity of the propagation of energy U is given by

$$U = \frac{\mu \frac{\partial u}{\partial t} \frac{\partial u}{\partial x}}{w}. \quad (5.45)$$

Thus, the velocity of energy propagation U does not necessarily coincide with the phase velocity c of the harmonic wave's motion. However, for simple harmonic motion if we substitute $u = A \cos(kx - \omega t)$ in equation (5.45) we obtain $U = \omega/k = c$, and the energy in this case propagates with the phase velocity.

5.4.1 Phase and group velocities

We have seen that, except for simple harmonic wave motion, the phase velocity $c = \omega/k$, is not, in general, equal to the energy-transport velocity U . For waves with more than one frequency, the phase velocity is a function of the frequency $c(\omega)$ or wave number $c(k)$. This implies that the wave number is a function of frequency $k(\omega)$ and vice versa for $\omega(k)$. In this case, as we will see in particular for surface waves in Chapter 12, we have the phenomenon of wave dispersion and can define the *group velocity* as

$$v = \frac{d\omega}{dk}. \quad (5.46)$$

This is the velocity of the propagation of the packets or groups of waves and we will see that it is the velocity corresponding to the frequencies that makes the phase stationary (Section 12.3). If we substitute $\omega = ck$ into equation (5.46), we obtain the relation between phase and group velocities,

$$v = c + k \frac{dc}{dk}. \quad (5.47)$$

Since energy is contained in wave packets, the group velocity coincides with the velocity of propagation of energy $v = U$. According to equation (5.47), if the phase velocity c depends on the wave number k , then the phase and group velocities are different, whereas if c is constant they are equal.

For the simple case of a monochromatic shear wave that propagates in the positive x direction,

$$u = A \cos(kx - \omega t). \quad (5.48)$$

The phase velocity is $c = \beta = \omega/k$. According to (5.41), the energy density is

$$w = \mu A^2 k^2 \sin^2(kx - \omega t). \quad (5.49)$$

The energy contained in one wave length λ can be obtained by integration:

$$W = \mu A^2 k^2 \int_0^\lambda \sin^2(kx - \omega t) dx, \quad (5.50)$$

and this results in

$$W = \pi \mu k A^2 = 2\pi^2 \beta \rho T \left(\frac{A}{T} \right)^2, \quad (5.51)$$

where we have taken into account that $k = 2\pi/(BT)$ and $\beta^2 = \mu/\rho$. This result can be applied also to P waves with a similar expression changing the velocity β into α . Hence, the energy contained in a wave length of a monochromatic wave is proportional to the square of the amplitude or the square of the amplitude divided by the period. This is used in the definition of the magnitude of an earthquake (Chapter 16).

We can also calculate the velocity of energy transport U using equation (5.45) by substitution of w from (5.49) and the derivatives of u from (5.48). The result is $U = \omega/k$; the velocity of energy propagation is equal, in this case, to the phase velocity. This result is consistent with the fact that the phase velocity is constant and therefore equal to the group velocity, according to (5.47).

In conclusion, in an elastic medium, energy is propagated by waves with a velocity that, in general, is equal to the group velocity and different from the phase velocity. In the case that the phase velocity is constant with frequency, the two velocities are equal.

5.5 The effect of gravity on wave propagation

The Earth possesses a *gravitational field* that affects the propagation of waves in its interior. In a non-rotational spherical Earth, the gravitational potential is $V = GM/r$. For points in its interior, the potential satisfies Poisson's equation:

$$\nabla^2 V = -4\pi G\rho. \quad (5.52)$$

With the passage of a wave, the potential at an element of volume is perturbed so that $\delta V = V - V_0$ is the difference between the perturbed and the unperturbed potentials. If the unperturbed and perturbed densities are ρ_0 and ρ , we have

$$\nabla^2(V - V_0) = -4\pi G(\rho - \rho_0). \quad (5.53)$$

In terms of the cubic dilation θ , this equation is

$$\nabla^2(V - V_0) = -4\pi G\rho\theta. \quad (5.54)$$

The force due to the change in the gravitational potential is $F = \rho\nabla(V - V_0)$. If we substitute this force into the equation of motion (4.71), we obtain

$$\alpha^2 \nabla \theta - \beta^2 \nabla \times \omega + \nabla(V - V_0) = \ddot{u}; \quad (5.55)$$

by taking the divergence in (5.55) and substituting (5.54) into it, we obtain

$$\alpha^2 \nabla^2 \theta + 4\pi G \rho \theta = \ddot{\theta}. \quad (5.56)$$

If we substitute as a solution $\theta = A \cos(kx - \omega t)$ and solve for $c = \omega/k$, we obtain

$$\begin{aligned} c &= \alpha(1 - \varepsilon)^{1/2} \\ \varepsilon &= \frac{4\pi G \rho}{\alpha^2 k^2} \end{aligned} \quad (5.57)$$

The phase velocity c is now a function of the wave number k and differs from α according to the value of ε . If, for the Earth's mantle (Chapter 11), we take as mean values $\alpha = 11 \text{ km s}^{-1}$ and $\rho = 4.4 \text{ g cm}^{-3}$, then, for waves of 1 s period ($k = 0.57 \text{ km}^{-1}$), $\varepsilon = 10^{-7}$ and the velocity of P waves is not affected by gravity. If we consider waves of 300 s ($k = 0.002 \text{ km}^{-1}$) period, the value of ε is 0.01 and the velocity will be decreased by approximately 1%. This shows that we need to consider the effect of gravity on wave propagation only for waves with very long periods.

5.6 Plane waves

The simplest geometry of the wave front is that of a plane and the corresponding waves are called *plane waves*. The equation of the wave front in Cartesian coordinates (named after René Descartes) is given by

$$S(x_1, x_2, x_3) = v_1 x_1 + v_2 x_2 + v_3 x_3.$$

According to equation (5.26), v_1 , v_2 , and v_3 are the direction cosines of the normal to the plane wave front which define the direction of propagation or *ray trajectory* or *ray path*.

In Cartesian coordinates the wave equation is

$$\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} = \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2}. \quad (5.58)$$

Solutions for monochromatic waves of frequency ω can be written as

$$f(x_1, x_2, x_3, t) = A \exp \{i[k(x_1 v_1 + x_2 v_2 + x_3 v_3) - \omega t + \varepsilon]\} \quad (5.59)$$

or

$$f(x_k, t) = A \exp [i(k_j x_j - \omega t + \varepsilon)], \quad (5.60)$$

where $k_j = k v_j$ is the *wave-number vector*. The equation of the wave front for a phase at a given time is

$$k_1 x_1 + k_2 x_2 + k_3 x_3 = \text{constant}. \quad (5.61)$$

If we are dealing with waves with plane wave fronts of infinite extension, it is implied that their source is at an infinite distance. This condition does not correspond to any real problem. However, if we are interested only in wave propagation and can assume that the source is at a very large distance, plane waves are a good simplifying approximation.

Plane-wave solutions for the potentials ϕ and ψ are

$$\phi = A \exp [ik_\alpha(v_j x_j - \alpha t + \varepsilon)], \quad (5.62)$$

$$\psi_k = B_k \exp [ik_\beta(v_j x_j - \beta t + \eta)], \quad (5.63)$$

where $k_\alpha = \omega/a$ and $k_\beta = \omega/\beta$ are the wave numbers of P and S waves. Displacements for P and S waves are obtained by substitution of (5.62) and (5.63) into (5.6) and (5.7):

$$u_k^P = A i k_\alpha (v_1, v_2, v_3) \exp [ik_\alpha(v_j x_j - \alpha t + \varepsilon)], \quad (5.64)$$

$$u_k^S = [(B_3 v_2 - B_2 v_3), (B_1 v_3 - B_3 v_1), (B_2 v_1 - B_1 v_2)] i k_\beta \exp [ik_\beta(v_j x_j - \beta t + \eta)]. \quad (5.65)$$

Cartesian components of amplitudes of displacements of P waves are proportional to the direction cosines of the rays corresponding to *longitudinal waves*. Displacements of S waves are perpendicular to the direction of propagation, which can be verified by taking the scalar product of $\mathbf{u}^S$ and $\mathbf{v}$, which is always zero. S waves are, therefore, *transverse waves*.

Another way of showing the properties of displacements of P and S waves is to consider the total elastic displacements in the form of plane waves (with the initial phase null):

$$u_i = C_i \exp [ik(v_j x_j - ct)], \quad (5.66)$$

where c is the phase velocity whose value we do not yet know. By substitution of (5.66) into the homogeneous equation of motion (4.74), we obtain three equations for the three components of the displacement that, after dividing by k^2 and the exponential, may be expressed in matrix form as

$$\begin{pmatrix} v_1^2 - \frac{\rho c^2 - \mu}{\lambda + \mu} & v_1 v_2 & v_1 v_3 \\ v_1 v_2 & v_2^2 - \frac{\rho c^2 - \mu}{\lambda + \mu} & v_2 v_3 \\ v_1 v_3 & v_2 v_3 & v_3^2 - \frac{\rho c^2 - \mu}{\lambda + \mu} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = 0. \quad (5.67)$$

This is a homogeneous system of equations for C_1 , C_2 , and C_3 , and the condition for the existence of a solution is that the determinant of the system is null. From this condition we obtain a cubic equation for c^2 that has two different solutions that correspond to the already known values of α and β (4.72) and (4.73). This result confirms what we knew from the solution for the potentials ϕ and ψ , namely that, in an infinite homogeneous, isotropic, elastic medium, there are only these two types of waves. The components of the displacement for each wave can be obtained by replacing the two values of c in equation (5.67).

As an example, if we have a wave that propagates in the positive direction of the x_1 axis, that is, $v_i = (1, 0, 0)$, equation (5.67) now has the simpler form,

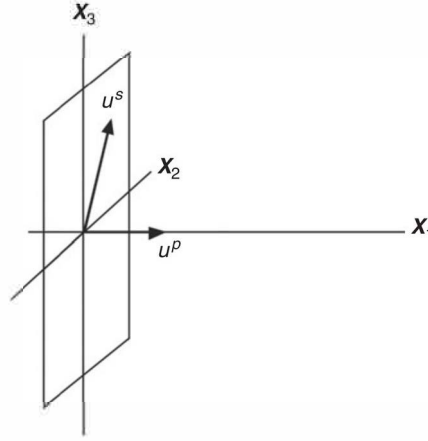


Fig. 5.4 Displacements of P and S plane waves propagating in the x_1 direction.

$$\begin{pmatrix} 1 - \frac{\rho c^2 - \mu}{\lambda} & 0 & 0 \\ 0 & \frac{\rho c^2 - \mu}{\lambda + \mu} & 0 \\ 0 & 0 & \frac{\rho c^2 - \mu}{\lambda + \mu} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = 0. \quad (5.68)$$

For the value $c = \alpha$, we obtain $C_1 \neq 0$ and $C_2 = C_3 = 0$. Displacements of the P waves are in the same direction as the propagation. For the value $c = \beta$, we obtain $C_1 = 0$, $C_2 \neq 0$, and $C_3 \neq 0$. Displacements of S waves are on the (x_2, x_3) plane, normal to the propagation direction (Fig. 5.4). The equations for displacements of the P and S waves are given by

$$u_k^P = (C_1, 0, 0) \exp[ik_\alpha(x_1 - \alpha t)], \quad (5.69)$$

$$u_k^S = (0, C_2, C_3) \exp[ik_\beta(x_1 - \beta t)]. \quad (5.70)$$

In the general case, displacements of P and S plane waves can be written as

$$u_k^P = D_k \exp[ik_\alpha(v_j x_j - \alpha t + \varepsilon)], \quad (5.71)$$

$$u_k^S = E_k \exp[ik_\beta(v_j x_j - \beta t + \eta)]. \quad (5.72)$$

Since displacements of P and S waves are perpendicular to each other, their scalar product must be null: $D_k E_k = 0$. The relations between the amplitudes of displacements (5.71) and (5.72) and those of potentials (5.62) and (5.63), according to (5.64) and (5.65), are

$$D_k = ik_\alpha v_k A, \quad (5.73)$$

$$E_k = ik_\beta e_{kjl} v_j B_l. \quad (5.74)$$

If the amplitudes of displacements are given in meters, those of potentials are given in m^2 . Their phases are shifted by $\pi/2$.

5.7 The geometry of P and S wave displacements

In seismology, the reference coordinate system generally used is the geographic one (x_1, x_2, x_3) , with x_1 and x_2 in the horizontal plane, x_1 positive to the north and x_2 positive to the west, and x_3 in the vertical direction positive up (zenith) (another choice is positive axes in the directions of north, east, and down (nadir)). The positive directions of the axes must satisfy a right-handed system. The direction of propagation of a wave is given by the vector $\mathbf{r}$ (v_1, v_2, v_3) and its projection onto the horizontal plane by $\mathbf{R}$. The angle between $\mathbf{r}$ and x_3 is the *angle of incidence* i ($0 - \pi/2$), and its complementary is $e = 90 - i$, the *angle of emergence*. The angle between $\mathbf{R}$ and x_1 is the *azimuth* a_z , measured from the north ($0 - 2\pi$) (Fig. 5.5). In terms of i and a_z , the direction cosines of the ray are

$$v_1 = \sin i \cos a_z, \quad (5.75)$$

$$v_2 = \sin i \sin a_z, \quad (5.76)$$

$$v_3 = \cos i. \quad (5.77)$$

In reference to the direction of wave propagation or the ray's direction given by the vector $\mathbf{r}$, we can define an orthogonal system of coordinate axes, the *propagation axes* $(\mathbf{R}, \mathbf{T}, \mathbf{Z})$; $\mathbf{R}$, already defined, is the horizontal component of the direction of the ray ($\mathbf{r}$), $\mathbf{T}$ is normal to $\mathbf{R}$ in the horizontal plane, and $\mathbf{Z}$ is vertical. $\mathbf{Z}$ is positive up, $\mathbf{R}$ is positive in the direction of propagation, and $\mathbf{T}$ is positive to the left, looking forward in the direction of propagation.

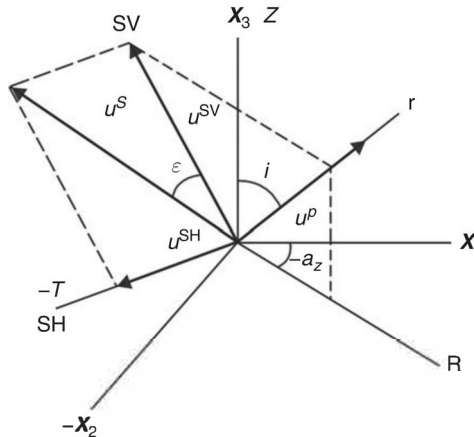


Fig. 5.5

Geometries of P and S wave displacements referred to geographic coordinate axes x_i (north, west, and zenith) and the propagation axes $(\mathbf{R}, \mathbf{T}, \mathbf{Z})$. Definitions of SH and SV components of the S wave are shown. a_z is the azimuth, i is the angle of incidence, and ϵ is the angle of polarization.

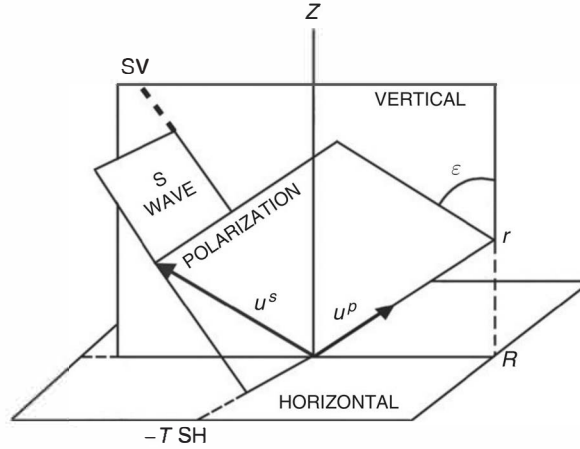


Fig. 5.6

Geometries of P and S wave displacements in reference to the four planes: vertical or incidence, horizontal, S wave motion, and S wave polarization.

Displacements of P waves have, in general, three components with respect to the geographic axes (u_1^P, u_2^P, u_3^P) and only two with respect to the propagation axes $(u_R^P, 0, u_Z^P)$. Displacements of S waves are contained on a plane normal to the vector $\mathbf{r}$ and have, in general, three components in both systems, (u_1^S, u_2^S, u_3^S) and (u_R^S, u_T^S, u_Z^S) .

In order to understand the geometry of plane P and S waves in Cartesian coordinates, it helps to define four planes: the *vertical plane* or *plane of incidence*; the *horizontal plane*; the *plane of S motion*; and the *polarization plane* of the S wave (Fig. 5.6). The plane of incidence contains the vectors $\mathbf{Z}$, $\mathbf{r}$, and $\mathbf{R}$. The plane of S motion is normal to the vector $\mathbf{r}$. The intersection of this plane with the horizontal plane defines the *SH* direction and that with the vertical plane defines the *SV* direction. In consequence, the displacement of the S wave can be divided into two components, namely, SH and SV. The polarization plane of the S wave contains the vectors $\mathbf{r}$ and $\mathbf{u}^S$. This plane forms an angle ε , called the *polarization angle*, with the incidence or vertical plane. In terms of the SH and SV components of S displacement, the angle of polarization is given by (Fig. 5.7c)

$$\tan \varepsilon = u_{SH}/u_{SV}. \quad (5.78)$$

The projection of the angle ε onto the horizontal plane is the angle γ , which is the angle between $\mathbf{R}$ and the horizontal component of S motion (u_H^S) (Fig. 5.7a):

$$\tan \varepsilon = \tan \gamma \cos i. \quad (5.79)$$

The unit vectors $\mathbf{r}$, $\mathbf{SH}$, and $\mathbf{SV}$ form another orthogonal system of axes. With respect to this system, amplitudes of P waves have components $(u_r, 0, 0)$ and those of S waves have $(0, u_{SH}, u_{SV})$. The relations among the different components of the amplitudes of P and S waves on the plane of incidence, plane of S motion, and horizontal plane are represented in Fig. 5.7.

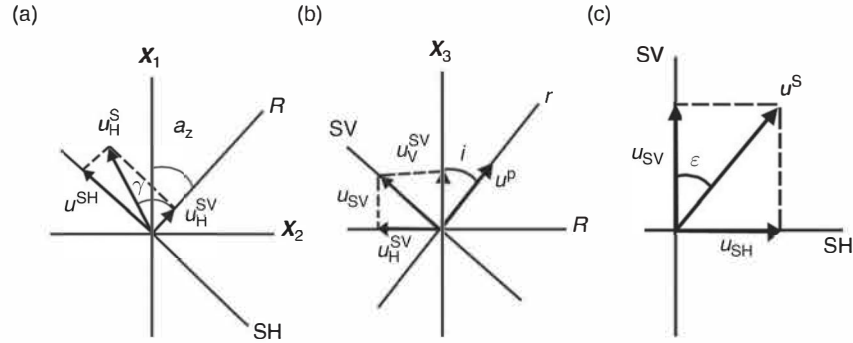


Fig. 5.7 Components of displacements of P and S waves: (a) on the horizontal plane, (b) on the vertical or incidence plane, and (c) on the plane of S wave motion. Planes are defined in Fig. 5.5 and Fig. 5.6.

5.8 Particular forms of the potentials

In many problems of propagation of plane waves in Cartesian coordinates, the relations between displacements and their potentials are simplified if the coordinate axes are selected to be in some particular direction. In general, expressions of the components of the displacements of P and S waves in terms of the scalar ϕ and vector ψ potentials are, according to (4.76),

$$u_1 = \frac{\partial \phi}{\partial x_1} + \frac{\partial \psi_3}{\partial x_2} - \frac{\partial \psi_2}{\partial x_3} = u_1^P + u_1^S, \quad (5.80)$$

$$u_2 = \frac{\partial \phi}{\partial x_2} + \frac{\partial \psi_1}{\partial x_3} - \frac{\partial \psi_3}{\partial x_1} = u_2^P + u_2^S, \quad (5.81)$$

$$u_3 = \frac{\partial \phi}{\partial x_3} + \frac{\partial \psi_2}{\partial x_1} - \frac{\partial \psi_1}{\partial x_2} = u_3^P + u_3^S. \quad (5.82)$$

If we take the axis x_1 as the direction of propagation, the displacements are

$$u_j = A_j \exp\{i[k(x_1 - ct) + \varepsilon]\}. \quad (5.83)$$

In this case, the displacements can be derived from three scalar potentials, ϕ , ψ , and Λ . The relations between the last two and the components of the vector potential ψ are $\psi = \psi_2$ and $\Lambda = -\psi_3$. The relations between the components of displacements and the scalar potentials are

$$u_1 = \frac{\partial \phi}{\partial x_1} = u^P, \quad (5.84)$$

$$u_2 = \frac{\partial \Lambda}{\partial x_1} = u^{SH}, \quad (5.85)$$

$$u_3 = \frac{\partial \psi}{\partial x_1} = u^{SV}. \quad (5.86)$$

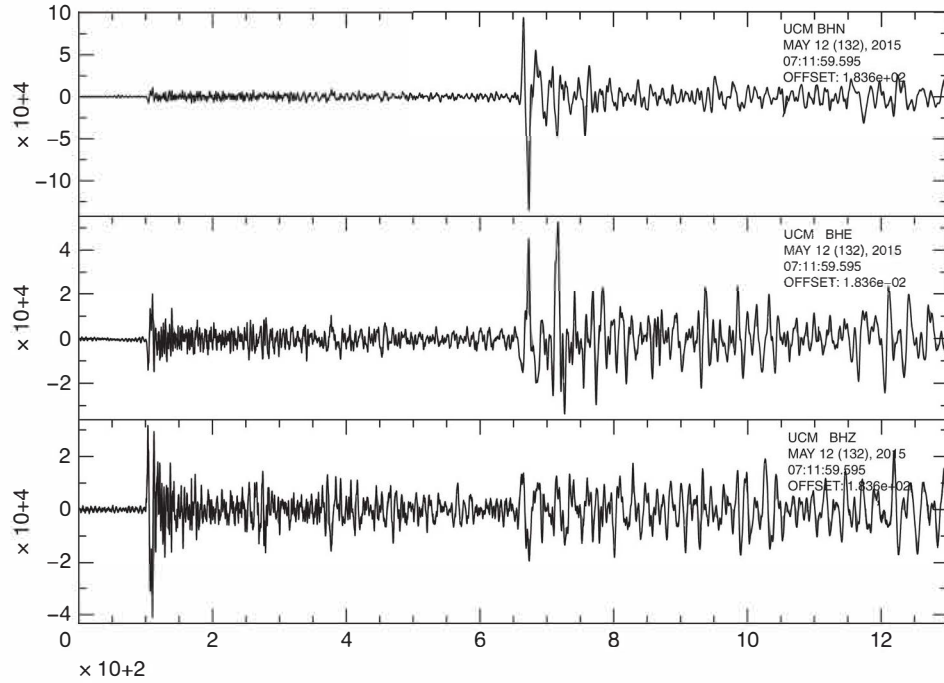


Fig. 5.8

P and S waves recorded at station UCM ($\Delta = 8012$ km) of the 2015 Nepal earthquake showing the arrivals of P and S waves. The P wave amplitude is largest at the vertical component (BHZ) because the ray is almost vertical with the incidence angle (i) of 17° (Fig. 5.5). Since the S wave lies on a plane normal to the ray (Fig. 5.6) its largest amplitudes are on the horizontal components (BHN and BHE). We can also see that the frequency of P waves is higher than that of S waves.

As we see, given the geometry of the problem, u_1 corresponds to the complete P wave and u_2 and u_3 to the SH and SV components of the S wave.

In many problems of wave propagation, it is convenient to use as the plane of incidence plane (x_1, x_3) , so that the ray is contained in such a plane (Fig. 5.9). The direction cosines of the ray $(v_1, 0, v_3)$ in terms of the angles of incidence i and emergence e are

$$v_1 = \sin i = \cos e, \quad (5.87)$$

$$v_3 = \cos i = \sin e. \quad (5.88)$$

In this case, the potentials and displacements are functions of x_1 and x_3 only. The components u_1 and u_3 can be derived from two scalar potentials ϕ and ψ , where $\psi = \psi_2$:

$$u_1 = \frac{\partial \phi}{\partial x_1} - \frac{\partial \psi}{\partial x_3} = u_1^P + u_1^{SV}, \quad (5.89)$$

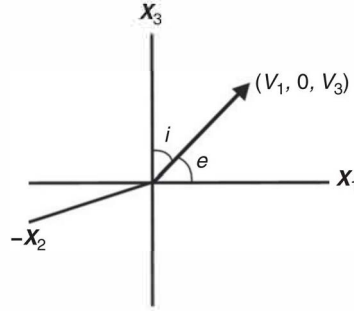


Fig. 5.9 A ray trajectory on the (x_1, x_3) plane, showing the angles of incidence i and emergence e .

$$u_3 = \frac{\partial \phi}{\partial x_3} + \frac{\partial \psi}{\partial x_1} = u_3^P + u_3^{SV}. \quad (5.90)$$

The u_2 component is usually kept separate, since, in this case, it is a solution of the wave equation (as in (5.8)). However, it can also be derived from a third scalar potential $\mathcal{A}$:

$$u_2 = -\frac{\partial \mathcal{A}}{\partial x_1} = u_2^{SH}. \quad (5.91)$$

The components u_1 and u_3 correspond to P and SV motion and the component u_2 corresponds to SH. The total displacement can be expressed in a single vector equation as a function of the ϕ , ψ , and $\mathcal{A}$ potentials:

$$\mathbf{u} = \nabla \phi + \nabla \times (0, \psi, 0) + \nabla \times (0, 0, \mathcal{A}). \quad (5.92)$$

This equation is a particular form of the more general relation that expresses displacements as functions of three scalar potentials (Eringen and Suhubi, 1974). We must remember that, when we use the vector potential ψ , it must satisfy the condition that its divergence is null. In consequence, only two of its components are independent and it is always possible to use three scalar potentials to represent the displacements.

5.9 Spherical waves

Earthquakes are generated at the source region by faulting from where seismic waves are propagated (Section 2.3), so the origin of the waves must be taken into account. For many practical purposes in seismology the source region can be approximated by a point from which waves propagate (Chapter 16). The problem of propagation of elastic waves generated at a point leads to the consideration of *spherical waves*. In this problem, we use spherical coordinates (r, θ, ϕ) and the components of displacements are u_r , u_θ , and u_ϕ (Appendix 2). If r is the direction of the ray, θ is measured from the vertical axis and ϕ on the horizontal plane, then u_r corresponds to the P wave, u_θ to SV, and u_ϕ to SH (Fig. 5.10).

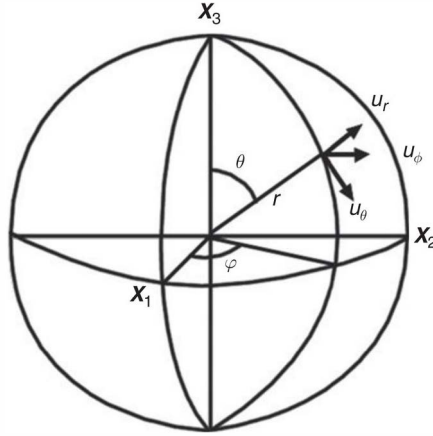


Fig. 5.10 Spherical coordinates (r, θ, ϕ) and components of the displacement (u_r, u_θ, u_ϕ) .

Components of displacements, as functions of potentials, are obtained by replacing the gradient and curl in spherical coordinates in equation (4.76) (Appendix 2):

$$u_r = \frac{\partial \Phi}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\psi_\phi \sin \theta) - \frac{1}{r \sin \theta} \frac{\partial \psi_\theta}{\partial \phi}, \quad (5.93)$$

$$u_\theta = \frac{1}{r} \frac{\partial \Phi}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial \psi_r}{\partial \phi} - \frac{1}{r} \frac{\partial}{\partial r} (r \psi_\phi), \quad (5.94)$$

$$u_\phi = \frac{1}{r \sin \theta} \frac{\partial \Phi}{\partial \phi} + \frac{1}{r} \frac{\partial}{\partial \theta} (r \psi_\theta) - \frac{1}{r} \frac{\partial \psi_r}{\partial \theta}. \quad (5.95)$$

The vector potential ψ must satisfy the condition of null divergence that, in spherical coordinates, is given by (A2.28).

For P waves that propagate from a point source with spherical symmetry and harmonic time dependence, we have only the potential $\Phi(r)$ and the Helmholtz equation is given by

$$\frac{d^2 \Phi}{dr^2} + \frac{2}{r} \frac{d\Phi}{dr} + k_a^2 \Phi = 0. \quad (5.96)$$

If we make the substitution $\Phi = f/r$, we obtain

$$\frac{d^2 f}{dr^2} + k_a^2 f = 0 \quad (5.97)$$

whose solution is

$$f(r) = A \exp(\pm i k_a r). \quad (5.98)$$

Adding the term of the harmonic time dependence for waves that propagate only in the positive direction of r , the potential $\Phi(r, t)$ is

$$\Phi(r, t) = \frac{A}{r} \exp[i(k_\alpha r - \omega t)]. \quad (5.99)$$

The amplitude of the potential decreases with distance by the factor $1/r$. If the energy generated at the source is finite, then, with increasing r , it must be distributed over an increasing area of the spherical wave and, in consequence, amplitudes per unit area decrease with distance by the factor of $1/r$. This decrease in amplitude with distance is called the *geometric spreading*. This term has not appeared for plane waves where amplitudes are constant with time, since their wave fronts are always infinite planes and their source is at an infinite distance, implying infinite energy.

If the problem has symmetry with respect to ϕ , displacements may be derived from three scalar potentials Φ , ψ , and A , which are functions of r , θ , and t :

$$u_r = \frac{\partial \Phi}{\partial r} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\psi \sin \theta), \quad (5.100)$$

$$u_\theta = \frac{1}{r} \frac{\partial \Phi}{\partial \theta} - \frac{1}{r} \frac{\partial}{\partial r} (r\psi), \quad (5.101)$$

$$u_\phi = -\frac{\partial A}{\partial \theta}. \quad (5.102)$$

Assuming a harmonic dependence on time, the potentials must satisfy the Helmholtz equation, which, for a potential $\Phi(r, \theta)$, is given by

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Phi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + k_\alpha^2 \Phi = 0. \quad (5.103)$$

Applying the method of separation of variables, $\Phi(r, \theta) = R(r)Y(\theta)$, we obtain the following differential equations:

$$r^2 \frac{d^2 R}{dr^2} + 2r \frac{dR}{dr} + [k_\alpha^2 r^2 - n(n+1)]R = 0, \quad (5.104)$$

$$\frac{1}{r^2} \frac{d^2 Y}{d\theta^2} + \cot \theta \frac{dY}{d\theta} + n(n+1)Y = 0, \quad (5.105)$$

where $n(n+1)$ is the constant of separation. By the substitution of $R = S/\sqrt{r}$ in equation (5.104), we obtain

$$\frac{d^2 S}{dr^2} + \frac{1}{r} \frac{dS}{dr} + \left(k_\alpha^2 - \frac{(n + \frac{1}{2})^2}{r^2} \right) S = 0. \quad (5.106)$$

If, in this equation, we substitute $x = k_\alpha r$, we obtain a differential *Bessel equation* (after Friedrich Wilhelm Bessel) whose solutions are *Bessel functions* of the type $J_{n+1/2}(k_\alpha r)$. These functions may be expressed in terms of *spherical Bessel functions* $j_n(k_\alpha r)$ (Appendix 3):

$$J_{n+1/2}(k_a r) = \left(\frac{2k_a r}{\pi} \right)^{1/2} j_n(k_a r). \quad (5.107)$$

For complex solutions, spherical *Hankel functions*, $h_n(z) = j_n(z) + in_n(z)$ (named after Hermann Hankel), are used (Appendix 3). Then, the solution for the function $R(r)$ is given by

$$R(r) = A_n \left(\frac{2k_a r}{\pi} \right)^{1/2} j_n(k_a r). \quad (5.108)$$

Making in equation (5.105) the change of variable $u = \cos \theta$, we obtain

$$(1 - u^2) \frac{d^2 Y}{du^2} - 2u \frac{dY}{du} + n(n+1) Y = 0. \quad (5.109)$$

This is a differential *Legendre equation* (named after Adrien-Marie Legendre) and its solutions are *Legendre functions*, $P_n(u)$ and $Q_n(u)$ (Appendix 3).

The complete solution for the potential $\Phi(r, \theta, t)$, adding the harmonic time dependence and using only Legendre functions of the first kind P_n and *Hankel spherical functions* h_n , is given by

$$\Phi(r, \theta, t) = A_n \left(\frac{2k_a r}{\pi} \right)^{1/2} h_n(k_a r) P_n(\cos \theta) e^{-i\omega t}. \quad (5.110)$$

For large values of r , using the asymptotic expression for $h_n(k_a r)$ in terms of harmonic functions, the solution may be written as

$$\Phi = A_n \frac{1}{r} \left(\frac{2}{k_a \pi} \right)^{1/2} P_n(\cos \theta) \exp \left[i \left(k_a r - \omega t - \frac{1}{2}(n+1)\pi \right) \right]. \quad (5.111)$$

For each value of n , there is a particular solution and the general solution is given by their sum. Solutions for the potentials ψ and A have the same form, with k_a replaced by k_β .

For the most general case in which there is no symmetry with respect to ϕ , displacements must be derived from scalar Φ and vector $(\psi_r, \psi_\theta, \psi_\phi)$ potentials that are functions of the three coordinates r, θ , and ϕ . For the scalar potential $\Phi(r, \theta, \phi)$, the Helmholtz equation is given by

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Phi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2} + k_a^2 \Phi = 0. \quad (5.112)$$

Using the method of separation of variables, $\Phi(r, \theta, \phi) = R(r)N(\theta)L(\phi)$, we obtain the following differential equations:

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \left(k_a^2 - \frac{n(n+1)}{r^2} \right) R = 0, \quad (5.113)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dN}{d\theta} \right) + \left(n(n+1) - \frac{m^2}{\sin^2 \theta} \right) N = 0, \quad (5.114)$$

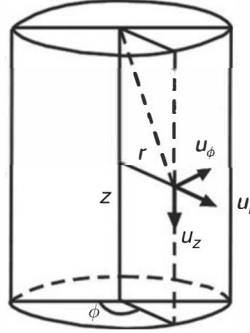


Fig. 5.11 Cylindrical coordinates (r, ϕ, z) and components of the displacement (u_r, u_ϕ, u_z) .

$$\frac{d^2 L}{d\phi^2} + m^2 L = 0, \quad (5.115)$$

where $n(n+1)$ and m^2 are the constants of separation. The solutions of equation (5.113), as in the previous case, may be expressed by Bessel or Hankel spherical functions. Solutions of equation (5.114) are *associate Legendre functions* (Appendix 3) and those of (5.115) are harmonic functions. The potential $\Phi(r, \theta, \phi, t)$ is given by the product of the three solutions and the harmonic function of time:

$$\Phi(r, \theta, \phi, t) = A_{nm} \left(\frac{2k_\alpha}{\pi} \right)^{1/2} h_n(k_\alpha r) P_n^m(\cos\theta) e^{i(\pm m\phi - \omega t)}. \quad (5.116)$$

As in the previous case, the general solution is the sum for all values of n . For each value of n , there are values of m from $m = -n$ to $m = n$. Then, for each value of n , there are $2n + 1$ solutions corresponding to different values of m . Solutions for the components of the vector potential $(\psi_r, \psi_\theta, \psi_\phi)$ are of similar form, with k_α replaced by k_β . The components of displacements u_r, u_θ , and u_ϕ can be obtained from the potentials by using equations (5.93)–(5.95).

5.10 Cylindrical waves

For many problems of wave propagation in seismology, for example, surface waves (Chapters 12–13) having axial symmetry, it is convenient to use cylindrical coordinates in their resolution in terms of *cylindrical waves*, using expressions for the gradient and curl in cylindrical coordinates (r, ϕ, z) (Appendix 2); expressions for the displacement components (u_r, u_ϕ, u_z) (Fig. 5.11) in terms of the scalar potential Φ and vector potential $(\psi_r, \psi_\phi, \psi_z)$ are

$$u_r = \frac{\partial \Phi}{\partial r} + \frac{1}{r} \frac{\partial \psi_z}{\partial \phi} - \frac{\partial \psi_\phi}{\partial z}, \quad (5.117)$$

$$u_\phi = \frac{1}{r} \frac{\partial \Phi}{\partial \phi} + \frac{\partial \psi_r}{\partial z} - \frac{\partial \psi_z}{\partial r}, \quad (5.118)$$

$$u_z = \frac{\partial \Phi}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} (r \psi_\phi) - \frac{1}{r} \frac{\partial \psi_r}{\partial \phi}. \quad (5.119)$$

The condition for null divergence of the vector potential $\boldsymbol{\psi}$, according to (A2.12), is

$$\frac{1}{r} \frac{\partial}{\partial r} (r \psi_r) + \frac{1}{r} \frac{\partial \psi_\phi}{\partial \phi} + \frac{\partial \psi_z}{\partial z} = 0. \quad (5.120)$$

In many problems we can assume symmetry with respect to ϕ and the equations are simplified. In this case, instead of the vector potential $\boldsymbol{\psi}$, we can use two scalar potentials, ψ and Λ , and the components of the displacements are

$$u_r = \frac{\partial \Phi}{\partial r} + \frac{\partial \psi}{\partial z}, \quad (5.121)$$

$$u_\phi = \frac{\partial \Lambda}{\partial \phi}, \quad (5.122)$$

$$u_z = \frac{\partial \Phi}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} (r \psi). \quad (5.123)$$

If waves propagate in the r direction and z is the vertical axis, u_r corresponds to P waves, u_z to the SV, and u_ϕ to the SH component of the S wave.

To study the solutions of the wave equation in cylindrical coordinates, we will consider first the case of axial symmetry (independence of ϕ). If we assume a harmonic dependence on time $\Phi(r, z)e^{i\omega t}$, the Helmholtz equation (5.10) for the potential $\Phi(r, z)$ is

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} + \frac{\partial^2 \Phi}{\partial z^2} + k_\alpha^2 \Phi = 0. \quad (5.124)$$

Using the method of separation of variables, $\Phi(r, z) = R(r) Z(z)$ (as in Section 5.2), the following equations are found:

$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + k_{\alpha r}^2 R = 0, \quad (5.125)$$

$$\frac{d^2 Z}{dz^2} + k_{\alpha z}^2 Z = 0, \quad (5.126)$$

where $k_\alpha^2 = k_{\alpha r}^2 + k_{\alpha z}^2$ and $k_{\alpha z}^2$ are the constants of separation. On making in equation (3.125) the change of variable $x = k_{\alpha r} r$, we obtain for $R(x)$ the equation

$$\frac{d^2 R}{dx^2} + \frac{1}{x} \frac{dR}{dx} + R = 0. \quad (5.127)$$

This is a differential Bessel equation for $n = 0$ and its solutions are the Bessel functions of zeroth order $J_0(k_{\alpha r} r)$ (Appendix 3). Equation (5.126) for the dependence of z has solutions

that are harmonic functions. Adding the harmonic time dependence, the complete solution of the potential Φ is

$$\Phi(r, z, t) = AJ_0(k_{\alpha}r)\exp[i(k_{\alpha}z - \omega t)]. \quad (5.128)$$

For large values of r , we can express Bessel functions in an asymptotic form in terms of harmonic functions:

$$\Phi(r, z, t) = \left(\frac{2}{\pi r k_{\alpha}}\right)^{\frac{1}{2}} A \exp\left[i\left(k_{\alpha}r + k_{\alpha}z - \omega t - \frac{\pi}{4}\right)\right]. \quad (5.129)$$

In this equation, we find the dependence of the amplitude of the potential on the inverse of the square root of the distance r . In this case, wave fronts are cylindrical surfaces whose areas increase with increasing r . Thus $1/\sqrt{r}$ is now the factor of geometric spreading due to the increase in area of the cylindrical wave front with increasing distance from the origin.

Solutions for the two other scalar potentials ψ and A are of similar forms to (5.128) and (5.129), with k_{α} replaced by k_{β} . Components of displacements are obtained from the three potentials Φ , ψ , and A using equations (5.121)–(5.123).

For the more general case in which the potentials depend on the three coordinates, for example, $\Phi(r, \phi, z)$, the Helmholtz equation is given by

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} + \frac{1}{r} \frac{\partial^2 \Phi}{\partial \phi^2} + \frac{\partial^2 \Phi}{\partial z^2} + k_{\alpha}^2 \Phi = 0. \quad (5.130)$$

Using the method of separation of variables, as before, $\Phi(r, \phi, z) = R(r)Y(\phi)Z(z)$, we obtain the following equations:

$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left(k_{\alpha}^2 - \frac{n^2}{r^2}\right)R = 0, \quad (5.131)$$

$$\frac{d^2 Y}{d\phi^2} + n^2 Y = 0, \quad (5.132)$$

$$\frac{d^2 Z}{dz^2} + k_{\alpha}^2 Z = 0. \quad (5.133)$$

The first equation is a differential Bessel equation and its solutions, finite at $r = 0$, are Bessel functions of order n . The other two equations (5.132) and (5.133) have harmonic functions for solutions. The solution for Φ , adding the harmonic time dependence, is

$$\Phi(r, \phi, z, t) = A_n J_n(k_{\alpha}r) \exp[i(k_{\alpha}z + n\phi - \omega t)]. \quad (5.134)$$

The solutions for components of the vector potential ψ have similar forms, replacing k_{α} by k_{β} . As in the previous case, components of displacements are obtained by using equations (5.121)–(5.123).

An important relation between the formulations in spherical coordinates (r, θ, ϕ) and cylindrical coordinates (ρ, z, ϕ) , when there is symmetry with respect to ϕ , is

$$\frac{1}{r} \exp(ik_r r) = \int_0^\infty J_0(k\rho) \exp(ik_z z) \frac{k}{k_z} dk, \quad (5.135)$$

where

$$\begin{aligned} r^2 &= \rho^2 + z^2 \\ k_r^2 &= k^2 - k_z^2 \end{aligned}$$

This equation, known as *Sommerfeld's relation* (named after Arnold Sommerfeld), allows one to relate problems of spherical and cylindrical symmetries.

5.11 Summary

Motion produced by earthquakes propagates through the Earth in the form of seismic waves. Their study is, then, a fundamental part of seismology. The equation of motion in an infinite homogeneous elastic medium can be easily transformed into two wave equations. Their solutions show that there are two types of waves: longitudinal or P waves and transverse or S waves, which propagate with velocities α and β respectively. Basic elements of waves in addition to their velocity are their amplitude and frequency or period, and their wave fronts and rays.

An important topic is the geometry of the displacements of the P and S waves and their relation with the ray path in terms of the geographical coordinates. The simplest geometry of wave fronts is that of plane waves with plane wave fronts. They are a good approximation when the source of waves is very far away. Waves propagating from a point source and with homogeneous conditions result in spherical waves with spherical wave fronts. In this case amplitudes decrease with the inverse of distance. Other cases, such as the propagation of surface waves, result in cylindrical waves with cylindrical wave fronts where amplitudes decrease with the inverse of the square root of distance.

5.12 Problems

P5.1 The elastic displacement potentials are

$$\begin{aligned} \phi &= 7 \exp \left[2i \left(\frac{1}{2} x_1 + \sqrt{3} x_2 / 2 - 6t \right) \right] \\ \psi_j &= (\sqrt{3} - 1, 6) \exp \left[3i \left(\frac{1}{2} x_1 + \sqrt{3} x_2 / 2 - 4t \right) \right]. \end{aligned}$$

Determine the amplitudes of the components of the P and S waves, and the angles i and a_z .

- P5.2 An S wave with only an SV component travels in the $\left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, \frac{\sqrt{3}}{2}\right)$ direction (x_3 is the vertical) and the modulus of its vector potential ψ is 5. If $k = 1$, find B_1 , B_2 , and B_3 , and u_1^S , u_2^S , and u_3^S .

- P 5.3 The displacement vector potential for S waves is

$$\psi_j = (-8\sqrt{3}, 2, 4) \exp \left[5i \left(\frac{1}{4}x_1 + \frac{\sqrt{3}}{2}x_2 + \frac{\sqrt{3}}{4}x_3 - 4t \right) \right].$$

If x_3 is the vertical, determine the SH and SV components, the direction cosines of the SH and SV directions, and the angle of polarization ε .

- P5.4 The amplitudes of displacement for the P and S waves are

$$\begin{aligned} u_i^P &= \left(4/(3\sqrt{2}), 4/(3/\sqrt{2}), 4\sqrt{3} \right) \\ u_i^S &= (-\sqrt{3}, -\sqrt{3}, \sqrt{2}) \end{aligned}$$

The velocities are $\alpha = 6 \text{ km s}^{-1}$ and $\beta = 4 \text{ km s}^{-1}$, and the frequency $\omega = 4$. Find the expressions for the potentials ϕ and ψ_i .

- P 5.5 Given the amplitudes of displacement of the P and S waves,

$$u_i^P = (4, 4, 8); \quad u_i^S = [8, 2\sqrt{2}, -(4 + \sqrt{2})],$$

find the components SV and SH and the angles i , a_z , ε , and γ .

- P5.6 Given the values $a_z = 60^\circ$, $\varepsilon = 30^\circ$, and $\gamma = 45^\circ$, if the modulus of the S wave is 5, find the components u_1^S , u_2^S , and u_3^S .

- P 5.7 Given the potentials ϕ with amplitude $A = 3$, and ψ with amplitudes $B_1 = -2$, $B_2 = 2$, and $B_3 = 0$, and the direction cosines of propagation $(\sqrt{2}/4, \sqrt{2}/4, \sqrt{3}/2)$, $k_a = \frac{2}{3}$, $T = \pi/2$, and $\sigma = \frac{1}{4}$, find the components of the displacements of P, SV, and SH waves.

- P 5.8 At the origin of an infinite medium with $\mu = 3 \times 10^{11} \text{ dyne cm}^{-2}$, $\rho = 3 \text{ g cm}^{-3}$, and $\sigma = \frac{1}{4}$, there is a center emitting plane waves. For a point at coordinates (500, 300, 141) km, write the expressions for the P and S cosine waves of arbitrary amplitudes. Find the arrival times.

- P 5.9 In an elastic medium of $\beta = 4 \text{ km s}^{-1}$ and $\sigma = \frac{1}{4}$, there are P and S waves of frequency 1 Hz propagating in the direction of $\left(1/\sqrt{3}, 1/\sqrt{3}, \sqrt{7}/(3\sqrt{2})\right)$. If the amplitude of potential ϕ is $A = \sqrt{3}/\pi$ and those of ψ are $B_1 = 3/\pi$ and $B_2 = \sqrt{2}/\pi$, find the amplitudes of the components of the displacements of P and S waves.

- P 5.10 In the Electronic Supplement (EM_5_P5.10) we have the recording of the N-S, E-W, and Z components of a P wave. The amplitude (vertical scale) is given in digital units (1DU = 10^{-9} m) and the time (horizontal scale) in seconds. Taking the seismograms to represent the wave displacements u^P :

- (a) Estimate the maximum amplitudes of the components of the displacements for the P wave.
- (b) Estimate the azimuth (a_z) and angle of incidence (i).
- (c) Write the equation for the scalar potential of the P wave.

The coefficient of Poisson σ is 0.25 and $\beta = 3\sqrt{3}$ km/s.

- P 5.11 The amplitude of S wave components (north, east, nadir) are (6, 3.23, 3). The azimuth is 60° and incidence angle 30° . Determine the SV and SH components.

In [Chapter 4](#) we considered the propagation of elastic waves in homogeneous media. We know that materials in the Earth are not homogeneous, rather their elastic properties vary with depth and from one region to another. This variation may be gradual, but there are also discontinuities that separate media with different densities and elastic coefficients. Let us consider now the phenomena that take place when waves propagate from one medium to another with different properties. When waves fall upon the surface separating the two media, part of the energy is *reflected* back into the first medium and part is *transmitted* or *refracted* into the second medium. Reflection of waves also occurs when there is a free surface. These problems are very important in seismology, since in the Earth there is a free surface and several discontinuities that separate media with different densities and elastic coefficients, as we saw in [Section 2.6](#). In some cases, such as for the crust and mantle, and for the mantle and core, the contrast between materials is large and produces notable phenomena of reflection and refraction of waves ([Chapter 11](#)). The theory of reflection and refraction of seismic waves was first developed by Cargill G. Knott in 1899 and Zöppritz in 1919. In this chapter, we will consider the problems of reflection and refraction between two media of different characteristics and at a free surface using plane geometry and waves.

6.1 Snell's law

The sine law that follows the trajectories of *incident*, *reflected*, and *refracted* waves was first proposed by Willebrord Snell in 1621 and further developed by Descartes in 1637 and Christiaan Huygens in 1678. In about 1650, Pierre de Fermat proposed his principle in mechanics and optics, that waves follow a trajectory for which the duration of the journey is stationary and minimum. From this principle, there follow two well-known consequences for the reflection and refraction of waves on a plane surface that separates two media. First, incident, reflected, and refracted rays are in the same plane, normal to the plane of separation of the two media, which is called the *plane of incidence*. Second, the trajectories of incident, reflected, and refracted rays follow Snell's law. For two media in which the waves' velocities are v and v' , if i is the angle between the incident ray in the medium of velocity v and the normal to the plane and i' is that of the refracted ray in the medium of velocity v' , this law establishes that

$$\frac{\sin i}{v} = \frac{\sin i'}{v'} = \frac{1}{c} = p, \quad (6.1)$$

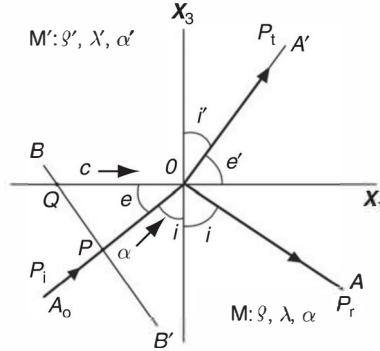


Fig. 6.1 Reflection (P_r) and refraction (P_t) of waves in two liquid media for waves incident from medium M (P_i). BB' wave front.

where p is called the *ray parameter* and $c = v/\sin i = v'/\sin i'$ is the component of the velocity in the direction parallel to the plane separating the two media. Also c is the apparent velocity of propagation of the points of intersection of wave fronts with the plane of separation. In Fig. 6.1, the point P travels a distance OP with velocity α during the same time as that in which the point Q travels a distance QO with velocity c . For reflection, the angles of incident and reflected rays are obviously the same, for both rays are in the same medium.

6.2 Reflection and refraction in two liquid media

We will start with the more simple case of the problem of the reflection and refraction of waves in two liquid media of densities ρ and ρ' and bulk moduli K and K' . In a liquid medium there are no shear waves ($\mu = 0$), only longitudinal or acoustic waves whose displacements are defined by only one potential, ϕ . Let us consider that the two media are separated by the plane (x_1, x_2) . The potentials ϕ and ϕ' in each medium are solutions of the wave equation (5.3) with velocities of propagation α and α' . For a liquid $\lambda = K$, and $\alpha^2 = K/\rho$. If we take the plane (x_1, x_3) as the incidence plane (azimuth $a_z = 0^\circ$) and express the direction cosines in terms of the *angle of incidence* i and *angle of refraction* i' (Fig. 6.1), then the potentials in each medium for plane waves of frequency ω are

$$\begin{aligned} \phi = & A_0 \exp [ik_\alpha(x_1 \sin i + x_3 \cos i - \alpha t)] \\ & + A \exp [ik_\alpha(x_1 \sin i - x_3 \cos i - \alpha t)], \end{aligned} \quad (6.2)$$

$$\phi' = A' \exp [ik_{\alpha'}(x_1 \sin i' + x_3 \cos i' - \alpha' t)], \quad (6.3)$$

where the wave numbers in each medium are

$$k_\alpha = \omega/\alpha, \quad (6.4)$$

$$k_{\alpha}' = \omega/\alpha'. \quad (6.5)$$

The potential ϕ represents the sum of the *incident* and *reflected* waves in the medium M and the potential ϕ' is for a wave *refracted* or *transmitted* in medium M' (Fig. 6.1). The amplitudes of the potentials of the incident, reflected, and refracted waves are A_0 , A , and A' . Equations (6.2) and (6.3) can be expressed in terms of the tangent of the angles of *emergence* ($e = \pi/2 - i$) in the form

$$\phi = A_0 \exp [ik(x_1 + x_3 \tan e - ct)] + A \exp [ik(x_1 - x_3 \tan e - ct)], \quad (6.6)$$

$$\phi' = A' \exp [ik(x_1 + x_3 \tan e' - ct)], \quad (6.7)$$

where, according to Snell's law,

$$k = k_{\alpha} \cos e = k_{\alpha}' \cos e', \quad (6.8)$$

$$c = \frac{\alpha}{\cos e} = \frac{\alpha'}{\cos e'}, \quad (6.9)$$

where k is the wave number associated with the apparent velocity c ($k = \omega/c$). From equation (6.9), we can derive

$$\tan e = (c^2/\alpha^2 - 1)^{1/2} = r, \quad (6.10)$$

$$\tan e' = (c^2/\alpha'^2 - 1)^{1/2} = r'. \quad (6.11)$$

Equations (6.6) and (6.7) can be expressed using (6.10) and (6.11) in terms of the velocities α , α' , and c . When there are reflected and refracted waves, c is larger than α and α' , and in consequence r and r' are real. Equations (6.6) and (6.7) for ϕ and ϕ' are of a very convenient form since, for $x_3 = 0$, all exponentials are equal.

The relations between the amplitude A_0 of the potential of the incident wave and A and A' of reflected and refracted waves are obtained by applying boundary conditions at the surface separating the two media ($x_3 = 0$). For two liquids, the boundary conditions are the continuity of the normal component of stress τ_{33} and of the displacement u_3 :

$$\tau_{33} = \tau'_{33}, \quad (6.12)$$

$$u_3 = u'_3. \quad (6.13)$$

For liquids, stress and displacement components can be expressed in terms of the potential ϕ . First we express the stress in terms of the strain (4.19), then the strain in terms of derivatives of displacements (4.20), taking into account that for a liquid $\mu = 0$, and, finally, displacements as functions of only the potential ϕ (4.85) ($u_i = \partial\phi/\partial x_i$), obtaining

$$\tau_{33} = \lambda(\phi_{,11} + \phi_{,33}) = \rho\omega^2\phi, \quad (6.14)$$

$$u_3 = \phi_{,3} = ik \tan e \phi, \quad (6.15)$$

where we have used the fact that, according to the wave equation, $\lambda \nabla^2 \phi = \rho \omega^2 \phi$. By substituting (6.14) and (6.15) into (6.12) and (6.13), with the potentials of each ray from equations (6.6) and (6.7), taking into account (6.10) and (6.11), we finally obtain for $x_3 = 0$,

$$A_0 + A = \frac{\rho'}{\rho} A', \quad (6.16)$$

$$A_0 - A = \frac{r'}{r} A'. \quad (6.17)$$

These two equations can be written in matrix form:

$$A_0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & \rho'/\rho \\ 1 & r'/r \end{bmatrix} \begin{bmatrix} A \\ A' \end{bmatrix}. \quad (6.18)$$

The *reflection coefficients* V and *refraction* or *transmission coefficients* W are defined as the quotients of reflected and refracted or transmitted amplitudes with respect to the incident amplitude:

$$V = \frac{A}{A_0} = \frac{\rho' r - \rho r'}{\rho' r + \rho r'}, \quad (6.19)$$

$$W = \frac{A'}{A_0} = \frac{2\rho r}{\rho' r + \rho r'}. \quad (6.20)$$

If, instead of using equations (6.6) and (6.7), we use (6.2) and (6.3), we obtain

$$V = \frac{\rho' \alpha' \cos i - \rho \alpha \cos i'}{\rho' \alpha' \cos i + \rho \alpha \cos i'}, \quad (6.21)$$

$$W = \frac{2\rho \alpha' \cos i}{\rho' \alpha' \cos i + \rho \alpha \cos i'}. \quad (6.22)$$

These two equations can also be expressed in terms of the *refractive index* $n = \alpha/\alpha'$, the *density contrast* $m = \rho'/\rho$ between the two media, and angles of emergence e and e' :

$$V = \frac{m \sin e - n \sin e'}{m \sin e + n \sin e'}, \quad (6.23)$$

$$W = \frac{2 \sin e}{m \sin e + n \sin e'}. \quad (6.24)$$

According to these equations, the relation between W and V is

$$W = \frac{1}{m} (1 + V). \quad (6.25)$$

The *mechanical impedance* Z is defined as the quotient relating the stress and particle velocity in a particular orientation. In the x_3 direction, for a liquid, according to (6.14) and (6.15), Z is given by

$$Z = \frac{\tau_{33}}{\dot{u}_3} = \frac{\rho\alpha}{\sin e}. \quad (6.26)$$

Equations (6.23) and (6.24) can be written in terms of Z as

$$V = \frac{Z' - Z}{Z' + Z}, \quad (6.27)$$

$$W = \frac{1}{m} \left(\frac{2Z'}{Z' + Z} \right). \quad (6.28)$$

If waves travel from a medium M' to medium M , the reflection and refraction coefficients V' and W' , according to (6.19) and (6.20), are

$$V' = \frac{A'}{A'_0} = \frac{\rho r' - \rho' r}{\rho r' + \rho' r}, \quad (6.29)$$

$$W' = \frac{A}{A'_0} = \frac{2r'\rho'}{\rho r' + \rho' r}. \quad (6.30)$$

The complete problem with waves traveling from both media in both directions is described in total by the four coefficients V , V' , W , and W' . The matrix formed by these four coefficients is called the *dispersion matrix*:

$$D = \begin{pmatrix} V & V' \\ W & W' \end{pmatrix}$$

The first column represents the partition of amplitudes between reflected and refracted waves for waves traveling from medium M to medium M' and the second represents the partition for waves traveling from M' to M . It is important to notice that V and W represent

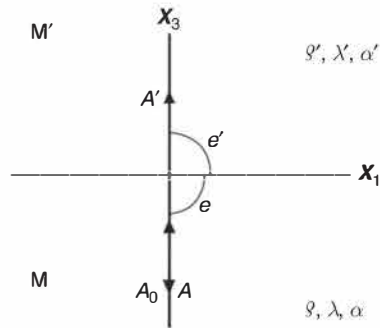


Fig. 6.2 Normal incidence in two liquids.

the partition of amplitudes of potentials but not the partition of energy between reflected and transmitted waves. For this reason their sum is not unity.

6.2.1 Normal incidence

A particular case of reflection and refraction corresponds to normal incidence, that is, when rays travel perpendicular to the surface separating the two media ($i = 0^\circ$, or $e = \pi/2$) (Fig. 6.2). The reflection and transmission coefficients are given by

$$V = \frac{\alpha' \rho' - \alpha \rho}{\alpha' \rho' + \alpha \rho}, \quad (6.31)$$

$$W = \frac{2\rho\alpha'}{\alpha' \rho' + \alpha \rho}. \quad (6.32)$$

We can see from these two equations that, if the contrast between the densities and velocities of the two media is small, V is very small and W is nearly unity. Waves are transmitted into the second medium with little reflection. For the opposite case, if the contrast is large, V is nearly unity and W is small. Waves are reflected back into the same medium with little transmission. The larger the contrast between the properties of the two media the larger the reflections. From equations (6.31) and (6.32) we can easily verify that the sum of V and W is not unity, for they do not represent the partition of energy between the reflected and refracted waves.

6.2.2 Critical incidence

According to Snell's law, if $\alpha' > \alpha$, there is an angle i_c or e_c , called the *critical angle* of incidence or emergence, which corresponds to rays refracted in the medium of velocity α' with angle $e' = 0^\circ$, or $i' = \pi/2$. Then, according to (6.1),

$$\frac{\sin i_c}{\alpha} = \frac{\cos e_c}{\alpha} = \frac{1}{\alpha'} = \frac{1}{c}, \quad (6.33)$$

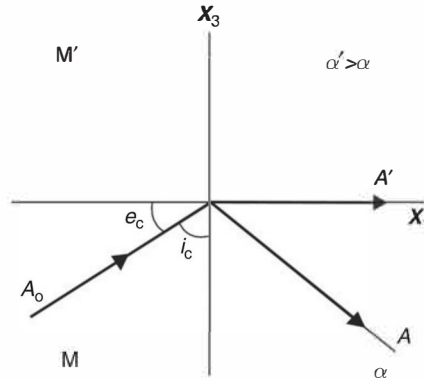


Fig. 6.3

Critical incidence in two liquids (e_c and i_c are the critical angles).

and, therefore,

$$\cos e_c = \sin i_c = \alpha/\alpha' = n. \quad (6.34)$$

If we substitute $e' = 0^\circ$ into equations (6.23) and (6.24) we obtain $V = 1$ and $W = 2\rho/\rho'$. Waves refracted in medium M' propagate along the surface separating the two media and therefore $c = \alpha'$ (Fig. 6.3). Such a wave is called a *critically refracted* or *head wave*. Its potential is given by

$$\phi' = 2A_0\rho/\rho' \exp [ik_\alpha'(x_1 - \alpha't)]. \quad (6.35)$$

Rays with angles of emergence $e < e_c$ or angles of incidence $i > i_c$ are not refracted into medium M' and their energy is completely reflected back into medium M . This situation is called *total reflection*. Thus, from the point of view of the angle of incidence, we have *subcritical* ($i < i_c$) and *supercritical* ($i > i_c$) reflections. Supercritical reflected waves carry more energy than do subcritical waves since, in the former case, no energy is transmitted to the second medium. For supercritical waves, according to (6.9) and (6.33), $\cos e'$ would have a value larger than unity, which cannot correspond to a real angle, and $c < \alpha'$. Hence, according to (6.11), $\tan e'$ is an imaginary number:

$$\tan e' = i(1 - c^2/\alpha'^2)^{1/2} = i\bar{r}'. \quad (6.36)$$

The reflection coefficient V (6.19) is now a complex number with modulus equal to unity:

$$V = \frac{\rho'r - i\rho\bar{r}'}{\rho'r + i\rho\bar{r}'} = e^{-i\theta}, \quad (6.37)$$

where

$$\tan\left(\frac{\theta}{2}\right) = \frac{\rho\bar{r}'}{\rho'r} = \frac{\rho(1 - c^2/\alpha'^2)^{1/2}}{\rho'(c^2/\alpha'^2 - 1)^{1/2}}. \quad (6.38)$$

On putting $A = A_0 V$ and replacing V from (6.37), the expression for the potential of supercritical reflected waves is given by

$$\phi = A_0 \exp [ik(x_1 - rx_3 - ct - \theta/k)]. \quad (6.39)$$

Supercritical reflected waves have the same amplitude as incident waves, but they have a phase shift of θ .

6.2.3 Inhomogeneous waves

As we have seen, waves with angles of incidence greater than the critical angle ($i > i_c$, or $e < e_c$) are not refracted into the second medium; all of the energy is reflected back. However, we can verify from (6.20) that W is not zero but rather has a complex value, since r' is imaginary. Therefore, there is a potential ϕ' that corresponds to some type of elastic perturbation in the medium M' . This perturbation is formed not by normal

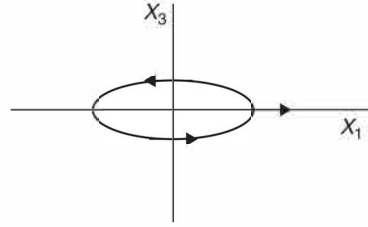


Fig. 6.4 Particle motion for an inhomogeneous wave produced by an incident wave with angle of incidence $e < e_c$.

transmitted waves but by waves of a special type. If we substitute the imaginary value of r' in expression (6.7) for ϕ' , we obtain

$$\phi' = A_0 W \exp[-k\bar{r}'x_3 + ik(x_1 - ct)], \quad (6.40)$$

where W is a complex number and $\bar{r}'$ is real. This potential represents waves that propagate in medium M' in the direction of x_1 with velocity c with values $\alpha < c < \alpha'$ and amplitudes decreasing exponentially with the distance x_3 from the surface separating the two media. This potential corresponds to an elastic perturbation in medium M' that exists only near the surface separating the two media. Owing to the special characteristics of these waves whose amplitudes decrease with distance x_3 along the same wave front, they are called *inhomogeneous* or *evanescent* waves.

Displacements of inhomogeneous waves also have different properties from those of normal transmitted waves. If we derive the displacements from the potential ϕ' in equation (6.40) according to (5.89) and (5.90) with ψ equal to zero, keeping only the real part, we obtain

$$u'_1 = -A_0 |W| k e^{-k\bar{r}'x_3} \sin[k(x_1 - ct) - \varepsilon], \quad (6.41)$$

$$u'_3 = -A_0 |W| k \bar{r}' e^{-k\bar{r}'x_3} \cos[k(x_1 - ct) - \varepsilon], \quad (6.42)$$

where the phase angle $\varepsilon = \tan^{-1}[\rho\bar{r}'/(\rho'r)]$.

Components u_1 and u_3 of displacements are shifted in phase by $\pi/2$ and in consequence the particles' motion is elliptical with its major axis in the direction of x_1 ($\bar{r}' < 1$) and of retrograde sense (Fig. 6.4). These are not, therefore, normal longitudinal waves of the type that exist in liquids, but waves of a special type with elliptical motion, similar in some ways to waves generated by wind on the surface of a liquid. Inhomogeneous waves are a phenomenon characteristic of supercritical reflections and we will see that they are related to surface waves in elastic media.

6.2.4 Reflected and transmitted energy

We have already mentioned that coefficients of reflection V and transmission W do not represent partition of energy and for this reason their sum is not unity. If we want to find the partition of the incident wave energy between reflected and refracted waves we have to

consider the energy that arrives at the surface of separation. For plane waves the energy contained in an element of volume $\delta V = \delta s \delta x$, where δs is an element of the wave front and δx is the distance traveled in the direction of propagation in an increment of time δt , according to (5.41), is the sum of kinetic and potential energies,

$$\delta E = \frac{1}{2} \left[\rho \left(\frac{\partial u}{\partial t} \right)^2 + \alpha^2 \rho \left(\frac{\partial u}{\partial x} \right)^2 \right] \delta s \delta x. \quad (6.43)$$

For a one-dimensional plane wave traveling in the x direction in a liquid of velocity α , $u = B \cos(kx - \omega t)$, and $\alpha = \omega/k$,

$$\frac{\partial u}{\partial x} = -\frac{1}{\alpha} \frac{\partial u}{\partial t}. \quad (6.44)$$

Since $\delta x = \alpha \delta t$, substituting into (6.43) gives

$$\delta E = \rho \left(\frac{\partial u}{\partial t} \right)^2 \alpha \delta s \delta t. \quad (6.45)$$

We define the *intensity* I as the energy per unit surface of the wave front and per unit time. From (6.45) the intensity is given by

$$I = \rho \alpha \left(\frac{\partial u}{\partial t} \right)^2. \quad (6.46)$$

If we derive the displacements from a potential ϕ ,

$$\phi = A \sin(kx - \omega t), \quad (6.47)$$

then, on substituting into (6.46), we can find the intensity in terms of the amplitude of the potential,

$$I = \alpha \rho k^2 \omega^2 A^2 = \frac{\rho \omega^4}{\alpha} A^2. \quad (6.48)$$

For normal incidence, an element of the wave front is equal to an element of the surface separating the two media and we can write directly the partition of energy on such a surface:

$$I_{\text{inc}} = I_{\text{refl}} + I_{\text{trans}}. \quad (6.49)$$

If we substitute values of intensities according to (6.48), knowing that the amplitudes of incident, reflected, and refracted potentials are, respectively, A_0 , VA_0 , and WA_0 , we obtain, after dividing by A_0 ,

$$1 = V^2 + \frac{\rho'}{\rho \alpha'} W^2. \quad (6.50)$$

This equation represents the partition of incident energy between reflected and refracted waves for normal incidence. If we substitute expressions for V and W from (6.31) and (6.32), we can easily verify this relation.

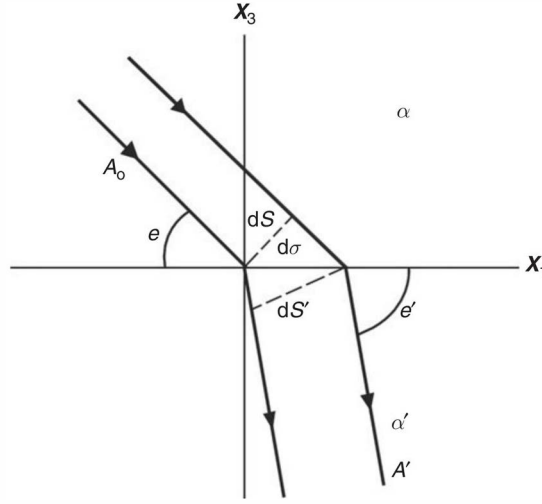


Fig. 6.5 The transmission of energy from a liquid medium to another by a beam of incident rays with element of wave-front area ds . The element of wave-front area of transmitted rays is ds' .

In the general case, we have to take into consideration the angle of incidence. Let us consider the energy contained in a ray bundle that occupies an element of the wave front of area ds , that strikes the surface separating the two media with an angle of emergence e (Fig. 6.5). The ray bundle of transmitted or refracted waves leaves the surface with an angle e' and occupies an element of the wave front of area ds' , while that of reflected waves occupies an element ds equal to that of incident waves. The element of area of the surface separating the two media $d\sigma$ occupied by the three ray bundles (incident, reflected, and refracted) must be the same. The relations among ds , ds' , and $d\sigma$ are

$$ds = d\sigma \sin e, \quad (6.51)$$

$$ds' = d\sigma \sin e'. \quad (6.52)$$

The energy incident on the element of area $d\sigma$ must be divided between the energies of the refracted and reflected waves:

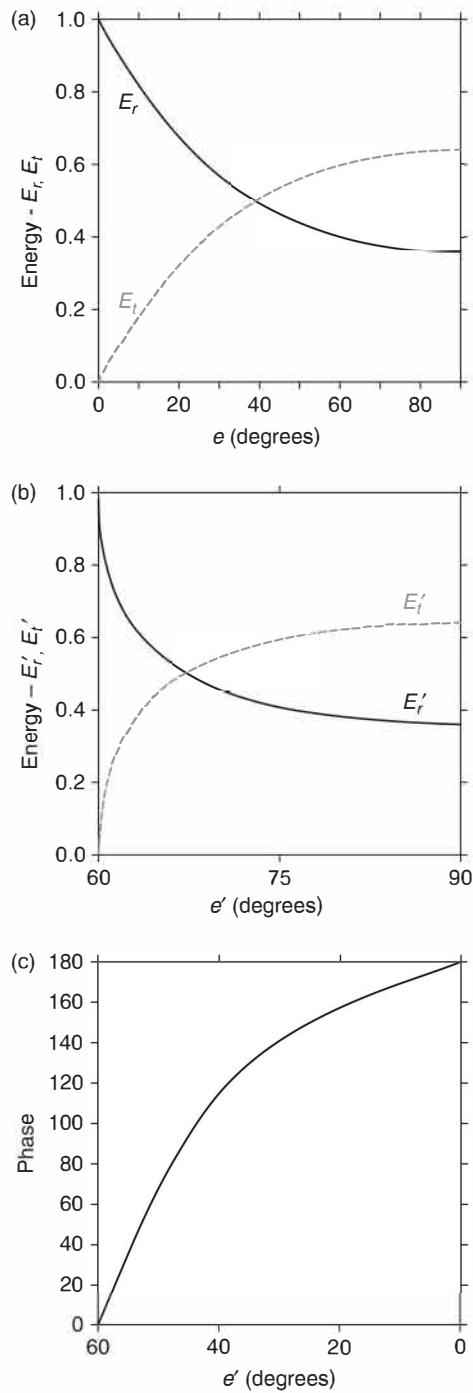
$$I_{\text{inc}} ds = I_{\text{refl}} ds + I_{\text{trans}} ds'. \quad (6.53)$$

On substituting equations (6.51) and (6.52) and intensities in terms of the amplitudes of incident, reflected, and refracted potentials (6.48), we finally obtain

$$1 = V^2 + \frac{\rho' \alpha \sin e'}{\rho \alpha' \sin e} W^2. \quad (6.54)$$

This equation represents the partition of incident energy between reflected and refracted waves for an arbitrary angle of incidence.

Figure 6.6 shows the partition of energy between reflected and refracted waves in two liquids with velocities $\alpha' = \alpha/2$ and densities $\rho' = \rho/2$ as a function of the angle of emergence.

**Fig. 6.6**

The partition of energy between reflected and transmitted waves in two liquids of velocities $a' = a/2$ and densities $\rho' = \rho/2$ as a function of the angle of emergence e . (a) Waves travel from medium M to medium M'. (b) Waves travel from medium M' to medium M; for $e' < 60^\circ$ all energy is reflected. (c) The phase shift between incident and reflected waves for angles of emergence less than the critical one ($e' < 60^\circ$).

In Fig. 6.6(a), waves travel from medium M to medium M'; this represents waves traveling from a medium of larger velocity to one of smaller velocity. For angles $e < 39^\circ$, the reflected energy is larger than the refracted energy, whereas for angles $e > 39^\circ$, the transmitted energy is larger. Figure 6.6(b) shows the case for waves traveling from medium M' to medium M, that is, from a medium of smaller velocity to one of larger velocity. The critical angle is $e'_c = 60^\circ$; for smaller angles waves are not transmitted and all the energy is reflected into medium M'. For angles larger than $e' = 68^\circ$, there is more energy transmitted than there is reflected. Figure 6.6(c) shows the phase shift of the reflected waves for supercritical incidence ($e' < e'_c$).

6.2.5 Reflection on a free surface

The case of the reflection of waves on a free surface (Fig. 6.7) is of special interest because of the presence of the free surface in the Earth. This is a particular case in which the second medium is the vacuum. The potential for incident and reflected waves is given by equation (6.6):

$$\phi = A_0 \exp [ik(x_1 + rx_3 - ct)] + A \exp [ik(x_1 - rx_3 - ct)]. \quad (6.55)$$

The boundary conditions on a free surface are that normal stresses or tractions are null. In our case, for $x_3 = 0$, $\tau_{33} = 0$. According to equation (6.14), this condition in terms of potentials leads to

$$\lambda(\phi_{,11} + \phi_{,33}) = 0. \quad (6.56)$$

By substitution of ϕ from (6.55), we obtain

$$A = -A_0. \quad (6.57)$$

The potential of reflected waves has the same amplitude as that of the incident wave but with the opposite sign; that is, there is a phase shift of π . The total displacement at the free surface is given by the sum of those of the incident and reflected waves, $U = u_{\text{inc}} + u_{\text{ref}}$. Using the potential ϕ from (6.55), taking derivatives according to (5.89) and (5.90) with ψ equal to zero, and using (6.57), we obtain for the real part that $U_1 = 0$ and U_3 is given by

$$U_3 = -2rkA_0 \sin [k(x_1 - ct)]. \quad (6.58)$$

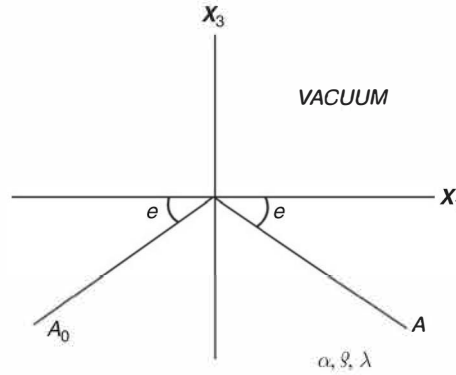


Fig. 6.7

Incident and reflected waves on a free surface of a liquid.

As we expected, the motion of the free surface of a liquid produced by the arrival of acoustic waves has only a vertical component and its amplitude depends on the angle of incidence of the waves.

6.3 Reflection and refraction in elastic media

Since in elastic media there are two types of waves, P and S, the problem of the reflection and refraction of waves is more complicated than that for two liquids. For incident P and S waves there are, generally, reflected and refracted P and S waves. This is due to Huygens' principle that states that, when waves reach a surface separating two media, its points become sources of waves that, in an elastic medium, are of both types. We take the plane (x_1, x_2) as that separating the two media and the plane (x_1, x_3) as the plane of incidence. The components of elastic displacements u_1 and u_3 can be derived from the scalar potentials ϕ and ψ , according to equations (5.89) and (5.90), and the component u_2 is kept apart. This way the potential ϕ represents P waves, whereas the potential ψ represents the SV and u_2 represents the SH components of S waves (Section 5.7).

If e and f are the angles of emergence of P and S incident and reflected waves in medium M of velocities α and β , and e' and f' are those of waves refracted or transmitted in medium M' with velocities α' and β' , then Snell's law (6.1) is given by

$$\frac{\cos e}{\alpha} = \frac{\cos f}{\beta} = \frac{\cos e'}{\alpha'} = \frac{\cos f'}{\beta'} = \frac{1}{c} = p, \quad (6.59)$$

where p and c have already been defined. As functions of the velocities, the tangents of e and f are

$$\tan e = (c^2/\alpha^2 - 1)^{1/2} = r, \quad (6.60)$$

$$\tan f = (c^2/\beta^2 - 1)^{1/2} = s, \quad (6.61)$$

and similarly for $\tan e' = r'$ and $\tan f' = s'$. The boundary conditions at the surface separating two elastic bodies ($x_3 = 0$) are continuity of the three components of the displacement u_i and continuity of the three components of stresses τ_{3i} across the surface (x_1, x_2) :

$$u_i = u'_i, \quad (6.62)$$

$$\tau_{3i} = \tau'_{3i}. \quad (6.63)$$

These conditions imply that the two media are welded together. We will treat separately the cases for incident SH (S waves with an SH component only), P, and SV waves.

6.3.1 Incident SH waves

Incident SH waves, owing to the selection of the orientation of the reference coordinate axes, have only the displacement component u_2 . It is easy to verify that, for only SH incident waves, the boundary conditions for u_1 , u_3 , τ_{31} , and τ_{33} result in homogeneous equations for the amplitudes of the reflected and refracted potentials ϕ , ψ , ϕ' , and ψ' , that

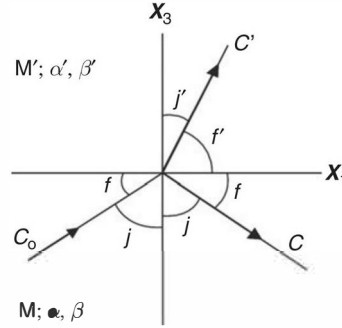


Fig. 6.8

Reflection and refraction in two elastic media for incident SH waves.

admit only a zero solution. This means that, if the incident S wave has only an SH component, then there are only reflected and refracted SH waves (Fig. 6.8). Thus, this case is similar to that of the two liquids we have already seen. The problem is reduced to that of the displacement component u_2 , which is given in both media by

$$u_2 = C_0 \exp [ik(x_1 + sx_3 - ct)] + C \exp [ik(x_1 - sx_3 - ct)], \quad (6.64)$$

$$u'_2 = C' \exp [ik(x_1 + s'x_3 - ct)], \quad (6.65)$$

where C_0 , C , and C' are the amplitudes of the incident, reflected, and refracted waves.

The boundary conditions at the surface separating the two media ($x_3 = 0$) are

$$u_2 = u'_2, \quad (6.66)$$

$$\tau_{32} = \tau'_{32}. \quad (6.67)$$

By expressing the stress τ_{32} in terms of the strain e_{32} (4.19), and this in terms of the derivatives of the displacement u_2 (4.20), and substituting the expressions (6.64) and (6.65) into the boundary conditions (6.66) and (6.67), we obtain the equations for the amplitudes in matrix form (just like in (6.18)):

$$C_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & \frac{s'\mu'}{s\mu} \end{pmatrix} \begin{pmatrix} C \\ C' \end{pmatrix}. \quad (6.68)$$

On solving for the amplitudes' quotients C/C_0 and C'/C_0 we obtain

$$V_{\text{SH}} = \frac{C}{C_0} = \frac{\mu s - \mu' s'}{\mu s + \mu' s'}, \quad (6.69)$$

$$W_{\text{SH}} = \frac{C'}{C_0} = \frac{2\mu s}{\mu s + \mu' s'}, \quad (6.70)$$

where V_{SH} is the reflection coefficient and W_{SH} is the refraction or transmission coefficient for SH waves. In this case, the coefficients are quotients of displacement amplitudes. Expressions (6.69) and (6.70) are similar to (6.19) and (6.20) obtained for the two liquids.

Here, also, the reflected energy is larger for a large contrast between the shear moduli μ and μ' of the two media.

If SH waves travel from medium M' into medium M , the reflection and transmission coefficients are

$$V'_{\text{SH}} = \frac{C'}{C_0} = \frac{\mu' s' - \mu s}{\mu' s' + \mu s}, \quad (6.71)$$

$$W'_{\text{SH}} = \frac{C}{C_0} = \frac{2\mu' s'}{\mu' s' + \mu s}. \quad (6.72)$$

The matrix for the total dispersion of SH motion between two elastic media is given by

$$\begin{pmatrix} V_{\text{SH}} & V'_{\text{SH}} \\ W_{\text{SH}} & W'_{\text{SH}} \end{pmatrix}.$$

In a similar way to what we did with the case of two liquids, we can find the partition of energy between the reflected and refracted waves for an arbitrary angle of emergence f :

$$1 = V_{\text{SH}}^2 + \frac{\rho' \beta \sin f'}{\rho \beta' \sin f} W_{\text{SH}}^2. \quad (6.73)$$

As in the case of two liquids, this equation shows the partition of the energy of the incident wave between the energy that is reflected back into the same medium and that which is transmitted into the second medium.

6.3.2 Critical incidence and inhomogeneous waves

When for the two media $\beta' > \beta$, then, for waves that travel from medium M to M' , there is a critical value for the angle of incidence, $f_c = \cos^{-1}(\beta/\beta')$, for which waves are refracted into medium M' with the angle $f' = 0$ and propagate along the surface separating the two media. For angles of incidence $f < f_c$, no waves are transmitted into the second medium and all the energy is reflected back into medium M (total reflection). For those values of f , s' is imaginary, since $c < \beta'$ and V_{SH} is a complex quantity of modulus unity. Reflected supercritical waves present a phase shift θ with respect to incident waves:

$$\tan\left(\frac{\theta}{2}\right) = \frac{\mu' \bar{s}'}{\mu s}, \quad (6.74)$$

where

$$\bar{s}' = (1 - c^2/\beta'^2)^{1/2}. \quad (6.75)$$

For supercritical reflections ($f < f_c$) we find also the phenomenon of the generation of inhomogeneous waves in medium M' , whose displacements are given by

$$u'_2 = C_0 W_{\text{SH}} \exp[-k \bar{s}' x_3 + ik(x_1 - ct)]. \quad (6.76)$$

These waves have the same characteristics as SH waves with only a u_2 component of displacement; they propagate in the x_1 direction with a velocity c ($\beta < c < \beta'$) that depends on f and their amplitude on the same wave front decreases exponentially with the distance (x_3) from the surface separating the two media. Note that as in the case of two liquids (Section 6.2.3) W_{SH} is a complex quantity.

6.3.3 Incident P and SV waves

For incident P waves, we have both reflected and refracted P and S waves (Fig. 6.9). Incident P waves that travel from medium M are represented by the potential ϕ with amplitude A_0 . Reflected P and S waves are represented by potentials ϕ (P) and ψ (SV) with amplitudes A and B , and a displacement component u_2 (SH) of amplitude C . Waves transmitted in medium M' are represented by potentials ϕ' (P) and ψ' (SV) with amplitudes A' and B' and a displacement component u'_2 (SH) with amplitude C' .

The relations between the amplitudes of incident P waves and those of reflected and refracted P and S waves are found by applying the boundary conditions of the continuity of stress and displacement at the surface separating the two media ($x_3 = 0$). From the conditions $u_2 = u'_2$ and $\tau_{32} = \tau'_{32}$ it is easily found that C and C' are zero. This means that, for an incident P wave, reflected and refracted S waves have only SV components.

The other four boundary conditions are

$$u_1 = u'_1, \quad (6.77)$$

$$u_3 = u'_3, \quad (6.78)$$

$$\tau_{31} = \tau'_{31}, \quad (6.79)$$

$$\tau_{33} = \tau'_{33}. \quad (6.80)$$

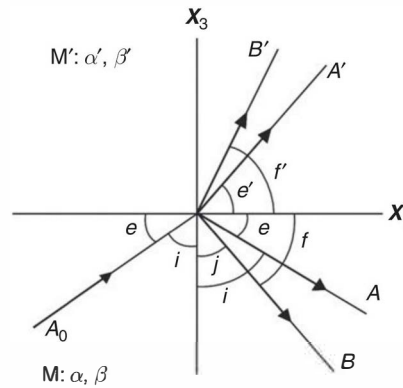


Fig. 6.9

Reflection and refraction in two elastic media for incident P waves. Reflected and refracted P and SV waves.

The potentials for P and SV motion in both media corresponding to incident, reflected, and refracted waves are

$$\phi = A_0 \exp [ik(x_1 + rx_3 - ct)] + A \exp [ik(x_1 - rx_3 - ct)], \quad (6.81)$$

$$\psi = B_0 \exp [ik(x_1 + sx_3 - ct)] + B \exp [ik(x_1 - rx_3 - ct)], \quad (6.82)$$

$$\phi' = A' \exp [ik(x_1 + r'x_3 - ct)], \quad (6.83)$$

$$\psi' = B' \exp [ik(x_1 + s'x_3 - ct)]. \quad (6.84)$$

On substituting in equations (6.77)–(6.80), and taking into account the relations between the stress and strain (4.19), strain and derivatives of displacements (4.20), and displacements and potentials (5.89) and (5.90), we obtain for the potentials the following equations:

$$\phi_{,1} - \psi_{,3} = \phi'_{,1} - \psi'_{,3}, \quad (6.85)$$

$$\phi_{,3} + \psi_{,1} = \phi'_{,3} + \psi'_{,1}, \quad (6.86)$$

$$\mu(2\phi_{,13} - \psi_{,33} + \psi_{,11}) = \mu'(2\phi'_{,13} - \psi'_{,33} + \psi'_{,11}), \quad (6.87)$$

$$\lambda(\phi_{,33} + \phi_{,11}) + 2\mu(\phi_{,33} + \psi_{,13}) = \lambda'(\phi'_{,33} + \phi'_{,11}) + 2\mu'(\phi'_{,33} + \psi'_{,13}). \quad (6.88)$$

By substitution into equations (6.85)–(6.88) of the expressions for the potentials (6.81)–(6.84) and writing the resulting equations in matrix form, we obtain

$$\begin{aligned} & A_0 \begin{pmatrix} 1 \\ r \\ 2r\mu \\ \lambda(1+r^2) + 2\mu \end{pmatrix} + B_0 \begin{pmatrix} -s \\ 1 \\ (1+s^2)\mu \\ 2\mu s \end{pmatrix} \\ &= \begin{pmatrix} -1 & -s & 1 & -s' \\ r & -1 & r' & 1 \\ 2r\mu & -(1-s^2)\mu & 2\mu' r' & \mu'(1-s'^2) \\ -\lambda(1+r^2) - 2\mu & -2\mu s & \lambda'(1+r'^2) + 2\mu' & 2\mu' s' \end{pmatrix} \begin{pmatrix} A \\ B \\ A' \\ B' \end{pmatrix}. \end{aligned} \quad (6.89)$$

For an incident P wave, $B_0 = 0$, and for an incident SV wave, $A_0 = 0$. The resulting expressions are known as the *Zöppritz equations*. For incident P waves, by solving the resulting equations from (6.89) with $B_0 = 0$, we determine the reflection and transmission coefficients. Now we have four coefficients corresponding to reflected and transmitted P and SV waves:

$$V_{PP} = \frac{A}{A_0}; \quad V_{PS} = \frac{B}{A_0}; \quad W_{PP} = \frac{A'}{A_0}; \quad W_{PS} = \frac{B'}{A_0}.$$

Reflected and transmitted SV waves from an incident P wave are often called *converted* waves.

For incident SV waves, the problem is similar. The reflection and transmission coefficients are derived from equation (6.89) by putting $A_0 = 0$. We have four coefficients and the converted reflected and transmitted waves are P waves:

$$V_{SS} = \frac{B}{B_0}; \quad V_{SP} = \frac{A}{B_0}; \quad W_{SS} = \frac{B'}{B_0}; \quad W_{SP} = \frac{A'}{B_0}.$$

Reflection and transmission coefficients for incident P and SV waves in isotropic media are functions of the elastic coefficients λ and μ , the densities ρ of the two media, and the angles of incidence e and f . The resulting expressions are simplified if we assume the value of Poisson's ratio $\sigma = 0.25$ ($\lambda = \mu$) (Section 4.2).

The dispersion matrix for the general P–SV motion generated in two elastic media in contact with incident waves traveling in both directions has 16 elements, eight for incident P and SV waves traveling from medium M and another eight for those traveling from medium M':

$$\begin{pmatrix} V_{PP} & V_{SP} & V'_{PP} & V'_{SP} \\ V_{PS} & V_{SS} & V'_{PS} & V'_{SS} \\ W_{PP} & W_{SP} & W'_{PP} & W'_{SP} \\ W_{PS} & W_{SS} & W'_{PS} & W'_{SS} \end{pmatrix}.$$

The partition of incident energy between the reflected and transmitted waves, P and SV, is deduced in a similar way to that which was presented for two liquids (6.54) and for the case of two solids for SH waves (6.73). For incident P waves we obtain

$$V_{PP}^2 + \frac{\alpha \sin f}{\beta \sin e} V_{PS}^2 + \frac{\rho' \alpha \sin e'}{\rho \alpha' \sin e} W_{PP}^2 + \frac{\rho' \alpha \sin f'}{\rho \alpha' \sin e} W_{PS}^2 = 1. \quad (6.90)$$

For incident SV waves,

$$V_{SS}^2 + \frac{\beta \sin e}{\alpha \sin f} V_{SP}^2 + \frac{\rho' \beta \sin f'}{\rho \beta' \sin f} W_{SS}^2 + \frac{\rho' \beta \sin e'}{\rho \alpha' \sin f} W_{SP}^2 = 1. \quad (6.91)$$

In both cases, the energy of an incident P or SV wave is divided into that of reflected and transmitted P and SV waves.

Just like incident SH waves, incident P and SV waves traveling from a medium of lesser velocity to one of greater velocity ($\alpha > \alpha'$) present the phenomenon of total reflection with critical angles given by $\cos e_c = \alpha / \alpha'$ and $\cos f_c = \beta / \beta'$. For supercritical angles we have also the generation of inhomogeneous waves. Since $\beta < \alpha$, there is also a critical angle for incident SV waves for which there is no reflected P wave: $\cos \bar{f}_c = \beta / \alpha$.

6.4 Reflection on a free surface

The problem of the reflection of elastic waves on a free surface is of particular interest in seismology since it represents the situation of the arrival of seismic waves at the Earth's

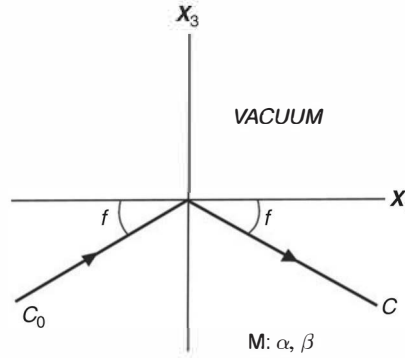


Fig. 6.10

SH waves incident on and reflected from a free surface of an elastic medium.

surface. The boundary conditions on a free surface, as we saw for a liquid, are that the components of stress across the surface are zero. We will treat first incident SH waves and then incident P and SV waves.

6.4.1 Incident SH waves

Let us consider an elastic half-space bounded by the free surface formed by the plane (x_1, x_2) . It is easy to show that, for incident SH waves, there are no reflected P or SV waves, but only reflected SH waves. The displacement u_2 of incident and reflected SH waves is given (Fig. 6.10) by

$$u_2 = C_0 \exp [ik(x_1 + sx_3 - ct)] + C \exp [ik(x_1 - sx_3 - ct)], \quad (6.92)$$

where C_0 is the amplitude of incident waves and C is that of reflected waves. On applying the boundary condition for $x_3 = 0$, we obtain

$$\tau_{32} = \mu u_{2,3} = 0. \quad (6.93)$$

On substituting (6.92) we obtain $C = C_0$ and that the reflection coefficient $V_{SH} = 1$. The reflected SH wave has the same amplitude as the incident wave.

6.4.2 Incident P waves

For an incident P wave there are reflected P and SV waves (Fig. 6.11). The nonexistence of reflected SH waves is also easily shown from the boundary condition $\tau_{32} = 0$. The potentials of incident P and reflected P and SV waves are

$$\phi = A_0 \exp [ik(x_1 + rx_3 - ct)] + A \exp [ik(x_1 - rx_3 - ct)], \quad (6.94)$$

$$\psi = B \exp [ik(x_1 - sx_3 - ct)]. \quad (6.95)$$

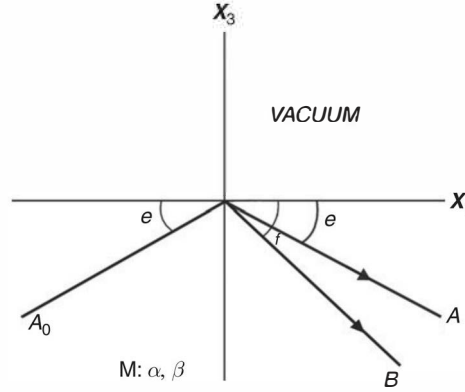


Fig. 6.11

Incident P and reflected P and SV waves for a free surface of an elastic medium.

The two pertinent boundary conditions at $x_3 = 0$ are that the stress components τ_{31} and τ_{33} are null. Writing the stress in terms of the strain for an isotropic material (4.19), the strain in terms of the derivatives of displacements (4.20), and, finally, displacements in terms of the potentials (5.89) and (5.90), we obtain

$$\tau_{31} = \mu(2\phi_{,13} - \psi_{,33} + \psi_{,11}) = 0, \quad (6.96)$$

$$\tau_{33} = \lambda\phi_{,11} + (\lambda + 2\mu)\phi_{,33} + 2\mu\psi_{,13} = 0. \quad (6.97)$$

For Poisson's relation ($\lambda = \mu$), by substituting the potentials (6.94) and (6.95) into equations (6.96) and (6.97), we obtain

$$(3r^2 + 1)(A_0 + A) - 2sB = 0, \quad (6.98)$$

$$2r(A_0 - A) + B(1 - s^2) = 0. \quad (6.99)$$

Using from Snell's law that $1 - s^2 = -(1 + 3r^2)$, we find that the coefficients of reflection are

$$V_{PP} = \frac{A}{A_0} = \frac{4rs - (1 + 3r^2)^2}{4rs + (1 + 3r^2)^2}, \quad (6.100)$$

$$V_{PS} = \frac{B}{A_0} = \frac{4r(1 + 3r^2)^2}{4rs + (1 + 3r^2)^2}. \quad (6.101)$$

These two coefficients represent the ratios of the amplitudes of the potentials of reflected P and SV waves with respect to that of incident P waves.

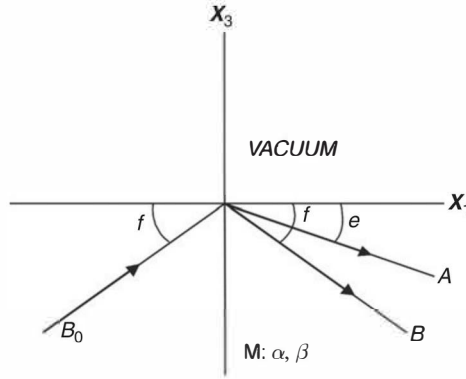


Fig. 6.12

Incident SV and reflected P and SV waves for a free surface of an elastic medium.

6.4.3 Incident SV waves

For incident SV waves, the problem is very similar. Again there are reflected SV and P waves, but no reflected SH waves (Fig. 6.12). The potentials of the incident and reflected waves are

$$\psi = B_0 \exp [ik(x_1 + sx_3 - ct)] + B \exp [ik(x_1 - sx_3 - ct)], \quad (6.102)$$

$$\phi = A \exp [ik(x_1 - rx_3 - ct)]. \quad (6.103)$$

By applying the boundary conditions as in the previous case, (6.96) and (6.97), and with the condition $\lambda = \mu$, we obtain

$$(1 - s^2)(B_0 + B) - 2rA = 0, \quad (6.104)$$

$$2s(B_0 - B) + (1 + 3r^2)A = 0. \quad (6.105)$$

The reflection coefficients are

$$V_{ss} = \frac{B}{B_0} = \frac{4rs - (1 + 3r^2)^2}{4rs + (1 + 3r^2)^2}, \quad (6.106)$$

$$V_{sp} = \frac{A}{B_0} = \frac{-4s - (1 + 3r^2)}{4rs + (1 + 3r^2)^2}. \quad (6.107)$$

Equations (6.98), (6.99), (6.104), and (6.105) can be expressed in matrix form by

$$\begin{pmatrix} 1 + 3r^2 \\ 2r \end{pmatrix} A_0 + \begin{pmatrix} 2s \\ 1 - s^2 \end{pmatrix} B_0 = \begin{pmatrix} -(1 + 3r^2) & 2s \\ 2r & -(1 - s^2) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (6.108)$$

From this equation, we can derive the particular cases for incident P or SV waves by letting $B_0 = 0$ or $A_0 = 0$.

6.4.4 Critical reflection of SV waves

According to Snell's law (6.59) for SV waves incident on a free surface, since $\alpha > \beta$, there exists a value of the angle of incidence f for which the angle e of reflected P waves is zero. This angle is called the critical angle f_c and is given by

$$\cos f_c = \beta/\alpha. \quad (6.109)$$

Incident SV waves with angles $f < f_c$ (supercritical SV waves) produce no reflected P waves, but only reflected SV waves. For these waves, we have $\beta < c < \alpha$ and therefore r is imaginary. In consequence, the reflection coefficient V_{ss} is a complex number:

$$V_{ss} = \frac{(1 - 3\bar{r}^2)^2 - i4\bar{r}s}{(1 - 3\bar{r}^2)^2 + i4\bar{r}s} = e^{-i\theta}, \quad (6.110)$$

where

$$\bar{r} = (1 - c^2/\alpha^2)^{1/2}, \quad (6.111)$$

$$\tan\left(\frac{\theta}{2}\right) = \frac{4\bar{r}s}{(1 - 3\bar{r}^2)^2}. \quad (6.112)$$

Reflected SV waves have the same amplitude as incident waves but with a phase shift θ . There are no reflected P waves, but there is a reflected potential ϕ that corresponds to an elastic perturbation in the form of inhomogeneous waves of P type that travel in the x_1 direction with velocity $c < \alpha$:

$$\phi = B_0 V_{sp} \exp [k\bar{r}x_3 + ik(x_1 - ct)]. \quad (6.113)$$

Since $r = i\bar{r}$ is imaginary, V_{sp} has complex values (6.107). This situation is similar to that found for supercritical incident waves in two elastic media when they travel from a medium of lesser velocity to one of larger velocity. Inhomogeneous P waves present in a half-space travel in the direction parallel to its free surface and their amplitudes are attenuated exponentially with the distance from this surface. This is an elastic perturbation that exists near the free surface and we will see that it is related to the generation of surface waves (Chapter 12).

6.4.5 The partition of energy

In the same way as for the reflection and refraction of waves in two elastic media, we can derive the partition of energy between P and SV waves reflected on a free surface. For incident P waves, the energy brought to the surface is divided between those of reflected P and SV waves (Fig. 6.13). In terms of energy intensities as in (6.53) we have the relation

$$I_{inc}^P ds = I_{refl}^P ds + I_{refl}^{SV} ds', \quad (6.114)$$

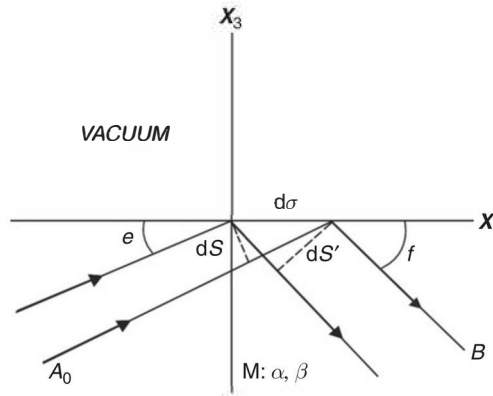


Fig. 6.13 An incident ray beam of P waves and a reflected ray beam for SV waves at a free surface of an elastic medium.

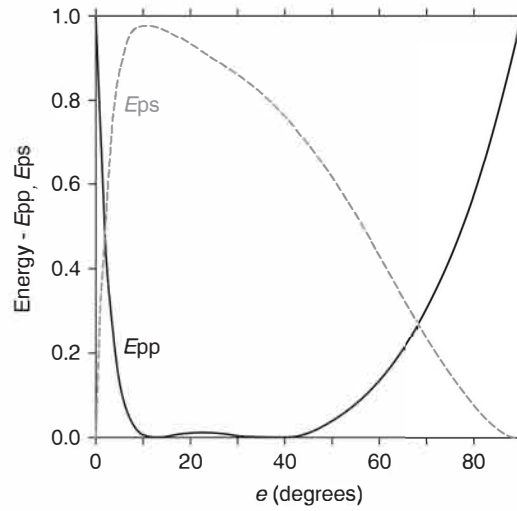


Fig. 6.14 The partition of energy between reflected P and SV waves for a P wave incident on a free surface of an isotropic elastic medium with $\lambda = \mu$.

where $ds/\sin e = ds'/\sin f = d\sigma$. On substituting ds and ds' in terms of $d\sigma$, the energy intensities as functions of the amplitudes of the potentials, as in (6.48), and those in terms of the coefficients of reflection for P and SV waves, the partition of energy is given by

$$V_{PP}^2 + \frac{\alpha \sin f}{\beta \sin e} V_{PS}^2 = 1. \quad (6.115)$$

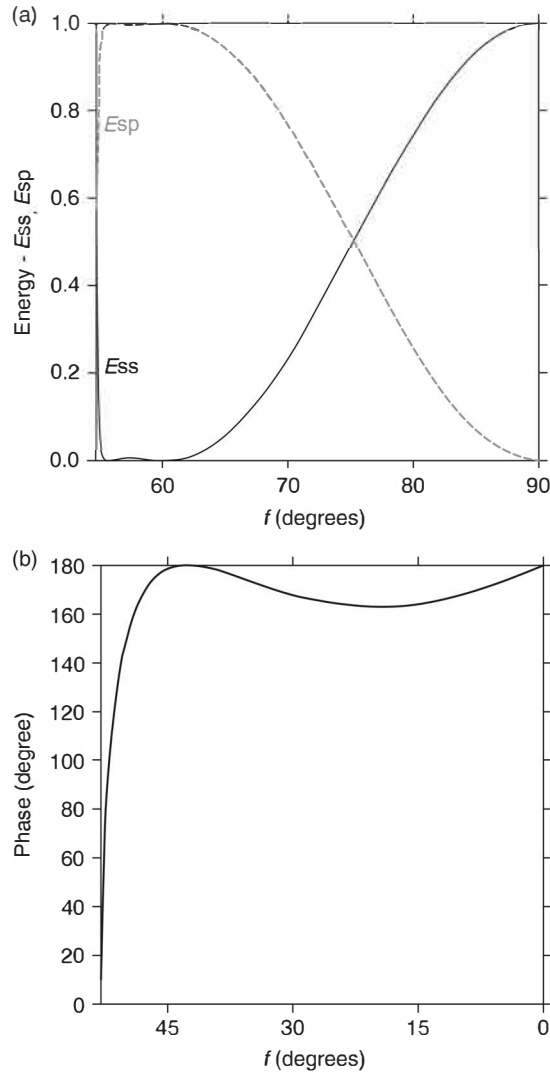


Fig. 6.15

(a) The partition of energy between reflected P and SV waves for an SV wave incident on a free surface of an isotropic elastic medium with $\lambda = \mu$; for $f < f_c$ there are no reflected P waves. (b) The phase shift between incident and reflected SV waves for supercritical angles of incidence ($f < f_c$).

For an incident SV wave, the relation is

$$V_{ss}^2 + \frac{\beta \sin e}{\alpha \sin f} V_{sp}^2 = 1. \quad (6.116)$$

Figures 6.14 and 6.15a show the partition of energy between reflected P and SV waves for incident P and SV waves with $\lambda = \mu$. For an incident P wave, $V_{pp} = 0$, for angles of incidence $e = 30^\circ$ and 12.47° ; and $V_{pp} = 1$ for $e = 0^\circ$ and 90° . For an incident SV wave, for angles of incidence less than the critical angle ($f_c = 54.73^\circ$), $V_{ss} = 1$ and there are no

reflected P waves. For normal incidence, the reflected waves are of the same kind as the incident ones. Figure 6.15b shows the phase shift of the supercritical reflected SV waves.

6.5 Motion at the free surface

Seismologic observations are carried out at the Earth's surface, thus seismograms are recordings of the total motion of this surface upon the arrival of seismic waves, not only of the incident waves (Chapter 3). Damage to buildings produced by earthquakes is also caused by the total motion of the Earth's surface. It is important, then, to study the total motion of the free surface of an elastic medium upon the arrival of waves that were traveling in its interior. The total displacement of points on a free surface is the sum of those of incident and reflected P and S waves. We will treat separately the motions of the surface upon the arrival of incident P and S waves.

6.5.1 Incident P waves

For an incident P wave, the total displacement of the free surface (U^P) is the sum of the displacements of the incident wave (u^{Pi}) and of the reflected P (u^{Pr}) and SV (u^{Sr}) waves (Fig. 6.16):

$$U^P = u^{Pi} + u^{Pr} + u^{Sr}. \quad (6.117)$$

If (x_3, x_1) is the plane of incidence, A_0 is the amplitude of the potential ϕ of incident P waves, and A and B are those of the potentials ϕ and ψ of reflected P and SV waves, (6.94) and (6.95); then the amplitudes of the horizontal and vertical components of the displacement at the free surface, according to (5.89) and (5.90), are

$$U_1^P = k(A_0 + A + B \tan f), \quad (6.118)$$

$$U_3^P = k(A_0 \tan e - A \tan e + B). \quad (6.119)$$

On substituting for the values of A and B in terms of A_0 and the reflection coefficients V_{PP} and V_{PS} , according to equations (6.100) and (6.101), we obtain

$$U_1^P = \frac{12A_0k}{D} \tan e \tan^2 f \sec e, \quad (6.120)$$

$$U_3^P = \frac{6A_0k}{D} (1 + 3 \tan^2 e) \tan e \sec^2 e, \quad (6.121)$$

$$D = 4 \tan e \tan f + (1 + 3 \tan^2 e)^2. \quad (6.122)$$

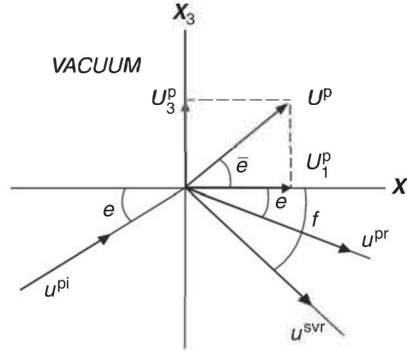


Fig. 6.16

The total motion of the free surface of an elastic medium for an incident P wave. The apparent angle of emergence is $\bar{e}$.

6.5.2 Incident S waves

For an incident S wave with SH and SV components, such that, according to the orientation of the coordinates axes, U_2^S is the SH component, and U_1^S and U_3^S are the horizontal and vertical components of SV motion (Fig. 6.17),

$$U_1^S = u_1^{Si} + u_1^{Sr} + u_1^{Pr}, \quad (6.123)$$

$$U_2^S = u_1^{Si} + u_2^{Sr}, \quad (6.124)$$

$$U_3^S = u_3^{Si} + u_3^{Sr} + u_3^{Pr}. \quad (6.125)$$

On substituting into these equations the amplitudes of SH displacement (C_0 incident and C reflected) and of the potential ψ of SV displacement (B_0 incident and B reflected) and A of the potential ϕ of reflected P waves, according to equations (6.92), (6.102), and (6.103), and the relations between the displacements and the potentials (5.89) and (5.90), we obtain

$$U_1^S = k(B_0 \tan f + A - B \tan f), \quad (6.126)$$

$$U_2^S = C_0 + C, \quad (6.127)$$

$$U_3^S = -k(-B_0 + A \tan e + B). \quad (6.128)$$

Just as in the previous case, using the relations between the reflected and incident amplitudes in terms of reflection coefficients, according to (6.106) and (6.107), and $C = C_0$, we obtain

$$U_1^S = \frac{6B_0 k}{D} (1 + 3 \tan^2 e) \tan f \sec^2 e, \quad (6.129)$$

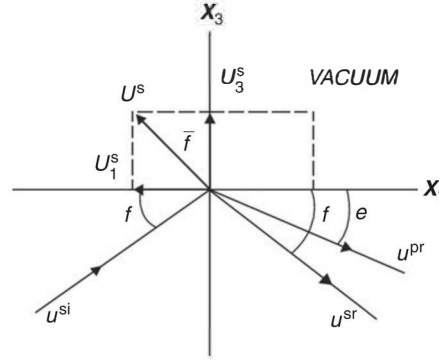


Fig. 6.17

The total motion of the free surface of an elastic medium for an incident SV wave. The apparent angle of emergence is $\bar{f}$.

$$U_2^S = 2C_0, \quad (6.130)$$

$$U_3^S = \frac{12B_0k}{D} \tan e \tan f \sec^2 e. \quad (6.131)$$

The amplitudes of the total displacement of the free surface depend on the amplitudes and angles of incidence of the incident waves. For incident P waves, the variables are A_0 and e ; for incident S waves, they are B_0 , C_0 , and f . Equations (6.120)–(6.122) and (6.129)–(6.131) allow also the determination of incident wave displacement amplitudes from the total motion observed at the free surface.

6.5.3 Apparent angles of incidence and polarization

Displacements observed at the Earth's surface correspond to the total motion of a free surface and thus include incident and reflected waves. If we want to calculate the incidence and polarization angles from amplitudes observed at the surface we have to take this fact into account. The angles of incidence $\bar{e}$ and $\bar{f}$ derived directly from the observed total motion for arriving P and S waves are called *apparent angles of incidence* (Figs. 6.16 and 6.17). If we substitute for incident P waves equations (6.120) and (6.121), and for incident S waves (6.129) and (6.131), we obtain

$$\tan \bar{e} = \frac{U_3^P}{U_1^P} = \frac{1 + 3 \tan^2 e}{2 \tan f}, \quad (6.132)$$

$$\tan \bar{f} = \frac{U_1^S}{U_3^S} = \frac{1 + 3 \tan^2 e}{2 \tan e}. \quad (6.133)$$

The S wave polarization angle γ (5.79) measured from total surface displacements is called the *apparent angle of polarization* $\bar{\gamma}$ and is found from the radial component, in our case U_1^S , and the transverse component, U_2^S of the observed motion at the surface for an arriving S wave (Fig. 6.18):

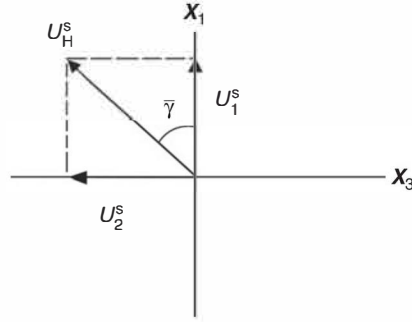


Fig. 6.18

The apparent polarization angle $\bar{\gamma}$ measured from the horizontal transverse (U_2) and radial (U_1) displacements at the free surface of an elastic medium for an incident S wave.

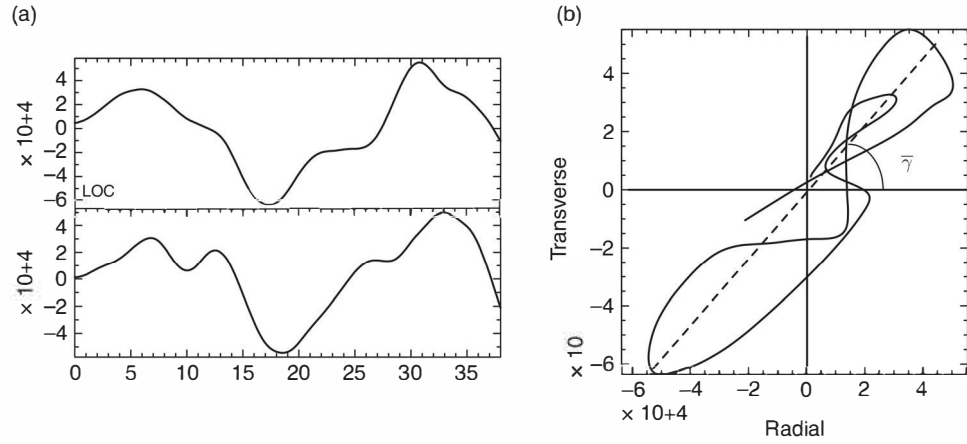


Fig. 6.19

(a) Radial (top) and transverse (bottom) components of S wave from seismograms of Figure 3.10. (b) Particle motion of S wave on the horizontal plane and apparent polarization angle $\bar{\gamma}$.

$$\tan \bar{\gamma} = \frac{U_2^S}{U_1^S} = \frac{C_0 [4 \tan e \tan f + (1 + 3 \tan^2 e)^2]}{3 B_0 k \tan f + (1 + 3 \tan^2 e) \sec^2 e}. \quad (6.134)$$

According to the definition of the polarization angle ε (5.78),

$$\tan \varepsilon = \frac{u_2^S}{(u_1^{S2} + u_3^{S2})^{1/2}} = \frac{C_0 \cos f}{B_0 k}. \quad (6.135)$$

On substituting (6.135) into (6.134) we obtain the relation between $\bar{\gamma}$ and ε :

$$\tan \bar{\gamma} = \frac{D}{3 \sin f \sec^2 e (1 + 3 \tan^2 e)} \tan \varepsilon, \quad (6.136)$$

where D is given in (6.122). By replacing equation (5.79) into (6.136) we obtain the relation between $\bar{\gamma}$ and γ . Equation (6.136) is valid only for angles of incidence less than the critical

one, that is, for $f > f_c$. When $f < f_c$, there are no reflected P waves, $\tan e$ is an imaginary quantity, and the factor in (6.136) is a complex quantity. There is a phase shift between the incident and reflected SV, but not for the SH waves, and the total motion of the free surface is not linear. Figure 6.19 shows the particle motion of the recorded horizontal components of S wave motion, obtained from the radial and transverse components of the seismogram, and the apparent polarization angle $\bar{\gamma}$ (Fig. 6.18).

In conclusion, the total motion of a free surface upon the arrival of P and S waves is complicated by the presence of reflected waves. For incident S waves and angles of incidence greater than the critical angle, the total motion is not linear. Thus, it is not always easy to find the amplitudes of incident waves.

6.6 Summary

The Earth is not a homogeneous medium but presents discontinuities in its composition with different elastic coefficients and density and has a free surface. It is, then, important to study the behavior of waves as they pass from one medium to another, as they are reflected into the same medium, and as they fall upon a free surface. Reflected and refracted rays between two media follow Snell's law (6.1). Incident energy is divided into that transmitted and reflected. The simpler case for liquid media where there is only one type of wave (P waves) has served as an introduction to the problem. The critical incidence is an important problem with the generation of inhomogeneous or head waves. For waves with incident angles greater than the critical, all energy is reflected.

In elastic media, where there are two types of waves, P and S, the situation is more complex. S waves must be separated into their SH and SV components and considered separately. For either P or SV incident waves there are both P and SV transmitted and reflected waves. The same happens when either P or SV waves fall upon a free surface. In each case there are both P and SV reflected waves. SH waves only generate SH reflected and refracted waves. Critical incidence has the added problem of incident SV waves for which there are no P reflected waves. Since seismological observations are at the Earth's free surface, motion recorded is the sum of that of incident and reflected waves.

6.7 Problems

- P6.1 For two liquid media in contact by the plane $x_3 = 0$, with parameters $\rho_1 = 1$, $K_1 = \frac{1}{2}$, $\rho_2 = \frac{3}{2}$, a P wave of frequency 3 Hz and potential amplitude 6 travels in the direction $(1/(2\sqrt{2}), 1/(2\sqrt{2}), \sqrt{3}/2)$ from medium 1 to medium 2. If the amplitude of the potential of the refracted wave is double that of the reflected one, find the expressions for the incident, refracted, and reflected potentials.

P6.2 A wave given by the potential

$$\phi = 4 \exp \left[\frac{1}{4} i \left(\frac{1}{\sqrt{6}} x_1 + \frac{1}{\sqrt{3}} x_2 + \frac{1}{\sqrt{2}} x_3 - 4t \right) \right]$$

travels from medium 1 to medium 2, which are both liquids, separated by the surface $x_3 = 0$. If the velocity of medium 2 is $\alpha_2 = 2 \text{ km s}^{-1}$, the pressure exerted by the incident wave on the surface is 5 dyn cm^{-2} , and the refracted energy is four times the reflected energy, find

- (a) the energy transmitted to the second medium;
- (b) the expressions for the refracted and reflected potentials.

P6.3 In two liquid media in contact, the velocities are $\alpha_1 = 4 \text{ km s}^{-1}$ and $\alpha_2 = 6 \text{ km s}^{-1}$, and $\rho_1 = 2 \text{ g cm}^{-3}$. For normal incidence from medium 1 to medium 2, the reflected potential amplitude is equal to the transmitted one. For a wave of frequency 1 Hz and potential amplitude $A_1 = 2000 \text{ cm}^2$, with angle of incidence $i_1 = 30^\circ$, the ratio of the transmitted to the reflected potential amplitude is $\frac{1}{3}$. For this wave, find the expression for the transmitted potential.

P6.4 Deduce the reflection and transmission coefficients A/A_0 , A'/A_0 , and B'/A_0 for a P wave incident from a liquid upon an elastic medium with $\sigma = \frac{1}{4}$.

P6.5 An incident wave in a semi-infinite elastic medium with Poisson ratio $\frac{1}{4}$ bounded by the surface $x_3 = 0$ is given by the potential

$$\phi = 10 \exp \left[5i \left(\frac{1}{2} x_1 + \frac{1}{2} x_2 + \frac{1}{\sqrt{2}} x_3 - 6t \right) \right].$$

Find the potential of the reflected S wave and the components of the displacement, u_1^S , u_2^S , and u_3^S .

P6.6 An incident S wave in a semi-infinite elastic medium of Poisson ratio $\frac{1}{4}$ bounded by the surface $x_3 = 0$ is given by the potential

$$\psi_j = (-10\sqrt{3}, 2, 4) \exp \left[5i \left(\frac{1}{4} x_1 + \frac{\sqrt{3}}{4} x_2 + \frac{\sqrt{3}}{2} x_3 - 4t \right) \right].$$

Find the components of the displacements of the reflected wave (u_1^S , u_2^S , u_3^S).

P6.7 An incident wave in a semi-infinite elastic medium with Poisson ratio $\frac{1}{4}$ bounded by the surface $x_3 = 0$ is given by the potential

$$\phi = 15 \exp \left[3i \left(\frac{1}{4} x_1 + \frac{\sqrt{3}}{2} x_2 + \frac{\sqrt{3}}{4} x_3 - 8t \right) \right].$$

Find the direction cosines and the total amplitude of the reflected S wave displacement.

- P6.8 An incident S wave in a semi-infinite elastic medium of Poisson ratio $\frac{3}{8}$ bounded by the surface $x_3 = 0$ is given by the potential

$$\psi_j = (2, -2, 0) \exp \left[i \left(\frac{1}{2\sqrt{2}}x_1 + \frac{1}{2\sqrt{2}}x_2 + \frac{\sqrt{3}}{2}x_3 - 4t \right) \right]$$

Find the SH and SV components of the reflected wave referred to the same system of coordinate axes as the incident wave.

- P6.9 An S wave represented by the potential

$$\psi_j = (-10\sqrt{3}, 2, 4) \exp \left[5i \left(\frac{1}{4}x_1 + \frac{\sqrt{3}}{4}x_2 + \frac{\sqrt{3}}{2}x_3 - 4t \right) \right]$$

is incident on the free surface ($x_3 = 0$) of an elastic half-space with $\sigma = \frac{1}{4}$. Find the components of the displacement of the reflected P wave.

- P6.10 An incident wave in a semi-infinite elastic medium with Poisson ratio $\frac{1}{4}$ bounded by the surface $x_3 = 0$ is given by the potential

$$\phi = 10 \exp \left[5i \left(\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{\sqrt{2}}x_3 - 6t \right) \right].$$

Find the components of the total displacement at the surface (U_i) and the apparent angle of incidence $\bar{e}$.

In the [previous chapter](#) we have seen the transmission and reflection of waves between media of different characteristics. This problem leads to the consideration of propagation of waves in stratified media, with waves being transmitted and reflected at each boundary between the layers. Many problems in seismology can be solved by representing the Earth as a stratified or layered medium, that is, one formed by layers of certain thicknesses and different mechanical properties. For certain problems we can use a flat approximation of parallel horizontal layers and thus the problems are thereby reduced to two dimensions. Layers of constant properties may also be considered as an approximation for media whose elastic coefficients vary in a continuous form with depth, which is a more realistic model for the Earth's interior. In layered or stratified media, problems are presented in discrete form and may be treated using matrix formulations. Solutions of problems of wave propagation in layered media using matrix formulation were introduced by William T. Thomson and Norman Haskell, becoming known as the *Thomson–Haskell method* (Thomson, 1950; Haskell, 1953). A similar formulation was proposed by Leon Knopoff (1964). George Gilbert and George Backus (1966) introduced the concept of the *propagator matrix* that allows a more generalized formulation of the problem (Kennett, 1983). This formulation is also very useful in the study of surface waves ([Chapter 12](#)).

7.1 Wave propagation in the (x, z) plane

Problems of wave propagation in two dimensions can be solved in a more convenient way if we use Cartesian coordinates, x horizontal and z vertical (positive downward) corresponding to the up to now used x_1 and x_3 . For a ray that forms an angle i (the angle of incidence) with the z axis ([Fig. 7.1](#)), the potential of P waves of a monochromatic harmonic wave of frequency ω can be expressed according to (6.2) as

$$\phi = A \exp [ik_{\bullet}(\sin i x + \cos i z - \alpha t)]. \quad (7.1)$$

Two other ways of writing this are

$$\phi = A \exp [i(k_x x + k_z z - \omega t)], \quad (7.2)$$

$$\phi = A \exp \left[i\omega \left(\frac{x}{c_x} + \frac{z}{c_z} - t \right) \right], \quad (7.3)$$

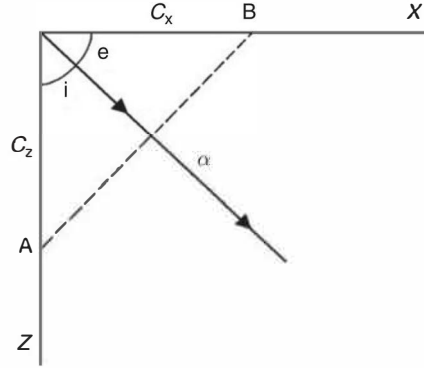


Fig. 7.1

A ray path or trajectory. The true velocity of propagation is α and the apparent velocities are c_x and c_z ; AB is the wave front.

where c_x and c_z are the *apparent velocities* of the wave front in the directions of the axes x and z (6.1), and k_x and k_z are the corresponding wave numbers. Thus, we have the relations

$$k_x^2 + k_z^2 = k_\alpha^2, \quad (7.4)$$

$$\frac{1}{c_x^2} + \frac{1}{c_z^2} = \frac{1}{\alpha^2}. \quad (7.5)$$

According to (7.1) and (7.3),

$$c_x = \frac{\alpha}{\sin i}; \quad c_z = \frac{\alpha}{\cos i}. \quad (7.6)$$

In many problems, for example when α is a function of z , it is convenient to separate the propagation in the z direction from that in the x direction. For this we call c and k the apparent velocity and wave number in the x direction ($c = c_x$). If we have waves propagating in positive and negative z directions (downward and upward), they can be expressed in the following way:

$$\phi = [A \exp(ikrz) + B \exp(-ikrz)] \exp[ik(x - ct)], \quad (7.7)$$

where $r = \tan e = \cot i$ (6.10), and also in the form

$$\phi = [A \exp(i\omega qz) + B \exp(-i\omega qz)] \exp[i\omega(px - t)], \quad (7.8)$$

where $p = 1/c = \sin i/\alpha$ and $q = 1/c_z = \cos i/\alpha$ are the inverses of the apparent velocities in the x and z directions, also called the *slowness* components in the two directions. As functions of c and α , r and q are given by

$$r = \left(\frac{c^2}{\alpha^2} - 1 \right)^{1/2}, \quad (7.9)$$

$$q = \frac{1}{c} \left(\frac{c^2}{\alpha^2} - 1 \right)^{1/2} = \left(\frac{1}{\alpha^2} - \frac{1}{c^2} \right)^{1/2} = pr. \quad (7.10)$$

In order to have waves propagating in the z direction, r and q must be real, or $c > \alpha$. If $c < \alpha$, r and q are imaginary, and, according to (7.7) and (7.8), waves propagate in the x direction and their amplitudes decrease or increase exponentially in the z direction. The exponential increase has no physical meaning and hence we must put $B = 0$ in (7.8). The exponential decrease corresponds to inhomogeneous waves, as we saw in the study of supercritical incidence (Sections 6.2 and 6.3). The same developments that have been applied to P waves can also be applied to S waves by replacing α by β . When the two types of waves are treated together, the following notation is used: for P waves, r, p_α, c_α , and q_α ; and for S waves, s, p_β, c_β , and q_β . Then, we have the relations $kr = \omega q_\alpha$ and $ks = \omega q_\beta$.

7.2 The equation for the displacement–stress vector

Let us now consider the propagation of S waves in a half-space limited by a free surface. For monochromatic plane waves of frequency ω that propagate in the plane (x, z) according to (7.7) and (7.8), one component of the displacement of S waves can be expressed in the following forms:

$$u = [A e^{iksz} + B e^{-iksz}] e^{ik(x-ct)}, \quad (7.11)$$

$$u = [A e^{i\omega qz} + B e^{-i\omega qz}] e^{i\omega(px-t)}, \quad (7.12)$$

where c, k, p, s , and q were defined in Section 6.3. In terms of the velocities, c and β , and angles of incidence i and f , s and q are given by equations (7.9) and (7.10) upon replacing α by β (Fig. 7.2):

$$s = \left(\frac{c^2}{\beta^2} - 1 \right)^{1/2} = \cot i = \tan f, \quad (7.13)$$

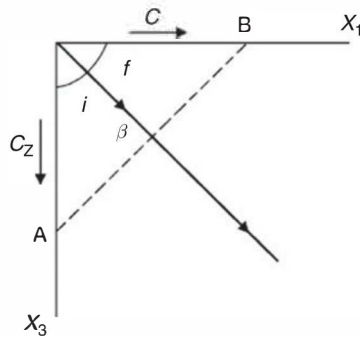


Fig. 7.2

The ray trajectory, and velocity components c in the x direction and c_z in the z direction. AB is the wave front.

$$q = \left(\frac{1}{\beta^2} - \frac{1}{c^2} \right)^{1/2} = \frac{\cos i}{\beta} = \frac{\sin f}{\beta}. \quad (7.14)$$

The relation between s and q is $ks = \omega q$.

In the general problem of an elastic medium, we have both P and S waves. Using expressions in the form of equations (7.11) and (7.12), in the exponential function we have s and q_β for S waves, and r and q_α for P waves (r and q_α are defined as in (7.13) and (7.14) by replacing β by α).

In the wave propagation corresponding to the propagation of motion there is also a propagation of stress. Components of the stress tensor across surfaces normal to z , that is, parallel to the boundary surfaces between layers, are τ_{3j} (subindex 1 in x direction, 3 in z , and 2 normal to the xz plane). Since rays propagate in the (x, z) plane, the problem can be separated into two parts: first, SH waves with displacements u_2 and stress component τ_{32} ; and second, P and SV waves with displacements u_1 and u_3 and stress components τ_{31} and τ_{33} . According to (7.11) and (7.12), the displacement and stress can be expressed in the forms:

$$u_i(z, x) = [u_1(z), u_2(z), u_3(z)] e^{ik(x-ct)}, \quad (7.15)$$

$$\tau_{3i}(z, x) = [\tau_{31}(z), \tau_{32}(z), \tau_{33}(z)] e^{ik(x-ct)}. \quad (7.16)$$

Since the term for propagation in the horizontal direction is the same in both equations, we can consider the displacement and stress as functions of z only. The displacement and stress can be put together, defining the *displacement–stress vector*. This is a useful concept because the boundary conditions at boundaries between layers involve continuity of both displacement and stress. For SH motion, it is given by the matrix

$$[u_2(z), \tau_{32}(z)], \quad (7.17)$$

and for P–SV motion it is given by

$$[u_1(z), u_3(z), \tau_{31}(z), \tau_{33}(z)]. \quad (7.18)$$

For SH motion in an isotropic homogeneous elastic medium for waves propagating in the (x, z) plane, the relation between the stress and the displacement is $\tau_{32} = \mu u_{2,3}$ (4.20), which can now be written as

$$\frac{du_2}{dz} = \frac{1}{\mu} \tau_{32}. \quad (7.19)$$

As we have already seen (Section 5.1), the transverse displacement $u_2(x, z, t)$ is a solution of the wave equation

$$\frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial z^2} = \frac{\rho}{\mu} \frac{\partial^2 u_2}{\partial t^2}. \quad (7.20)$$

By substituting (7.11) into (7.20), considering that $\tau_{32,3} = \mu u_{2,33}$, and dividing by the exponential, we obtain the relation

$$\frac{d\tau_{32}}{dz} = (\mu k^2 - \rho \omega^2) u_2 = k^2 \rho (\beta^2 - c^2) u_2. \quad (7.21)$$

Equations (7.19) and (7.21) can be expressed as a single matrix equation for the displacement–stress vector of SH motion (7.17):

$$\frac{d}{dz} \begin{pmatrix} u_2 \\ \tau_{32} \end{pmatrix} = \begin{pmatrix} 0 & 1/\mu \\ k^2 \rho (\beta^2 - c^2) & 0 \end{pmatrix} \begin{pmatrix} u_2 \\ \tau_{32} \end{pmatrix}. \quad (7.22)$$

For P–SV motion, the derivation of an equation similar to (7.22) is more complicated. The displacement components u_1 and u_3 are not solutions of the wave equation. These must be expressed in terms of scalar potentials ϕ and ψ according to (5.84) and (5.86). The components of the stress τ_{31} and τ_{33} also must be written in terms of ϕ and ψ , as in Section 6.3.3. Finally, the matrix equation for the displacement–stress vector is (Kennett, 1983)

$$\frac{d}{dz} \begin{pmatrix} u_1 \\ u_3 \\ \tau_{31} \\ \tau_{33} \end{pmatrix} = \begin{pmatrix} 0 & k^2 & (\rho \beta^2)^{-1} & 0 \\ 1 - 2\beta^2/\alpha^2 & 0 & 0 & (\rho \alpha^2)^{-1} \\ \omega^2 \rho h & 0 & 0 & k^2 (1 - 2\beta^2/\alpha^2) \\ 0 & -\rho \omega^2 & -1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_3 \\ \tau_{31} \\ \tau_{33} \end{pmatrix} \quad (7.23)$$

$$h = 4\beta^2 \frac{1}{c^2} \left(1 - \frac{\beta^2}{\alpha^2} \right) - 1.$$

The matrix equations (7.22) and (7.23) for SH and P–SV motion show that the derivatives with respect to z of the components of the displacement–stress vector are linear functions of the same components. The matrix of proportionality coefficients contains the elastic parameters of the medium (α , β , and ρ), the frequency ω , and the horizontal velocity c . Equations (7.22) and (7.23) are fundamental for the development of matrix methods for the resolution of wave propagation in layered media.

7.3 The propagator matrix

Equations (7.22) and (7.23) can be written in a general matrix form as

$$\frac{d\mathbf{b}(z)}{dz} = \mathbf{A}(c, z) \mathbf{b}(z), \quad (7.24)$$

where $\mathbf{b}$ is the displacement–stress vector defined in (7.17) and (7.18) and $\mathbf{A}$ is a square matrix whose elements are the elastic parameters and density of the medium as functions of the depth ($\alpha(z)$, $\beta(z)$, and $\rho(z)$), frequency ω , and horizontal velocity c . In the case of layers of constant parameters, in each layer, the only variables are c and ω .

For equations of the type (7.24), a matrix $\mathbf{B}$ that is a solution of the same equation is called a *fundamental matrix* of the equation (Gilbert and Backus, 1966). Thus we can write

$$\frac{d\mathbf{B}}{dz} = \mathbf{A}\mathbf{B}. \quad (7.25)$$

For a fixed value of $z = z_0$, a reference depth, the vector $\mathbf{b}(z_0)$ is a solution of (7.24). For the same depth, there exists also a solution of (7.25) that is the constant matrix $\mathbf{B}(z_0)$. We can normalize the problem and make $\mathbf{B}(z_0) = \mathbf{I}$, the unit matrix. We define now a new matrix $\mathbf{P}(z, z_0)$ that is a solution of equation (7.25) and thus a fundamental matrix:

$$\mathbf{P}(z, z_0) = \mathbf{B}^{-1}(z_0)\mathbf{B}(z). \quad (7.26)$$

Then, it follows that, for $z = z_0$, $\mathbf{P}(z_0, z_0) = \mathbf{I}$. If z is an arbitrary level, for an intermediate level z_1 between z_0 and z , according to (7.26), we can write

$$\mathbf{P}(z, z_0) = \mathbf{P}(z, z_1)\mathbf{P}(z_1, z_0). \quad (7.27)$$

If z_n is the n th level starting from z_0 , we can write the matrix $\mathbf{P}(z_n, z_0)$, using (7.27), which relates the n th level to the zeroth level, as a product of the matrices that relate all the intermediate levels step by step:

$$\mathbf{P}(z_n, z_0) = \prod_{i=1}^n \mathbf{P}(z_i, z_{i-1}). \quad (7.28)$$

Considering the properties of the matrix $\mathbf{P}(z, z_0)$, we have that, for the zeroth level, $\mathbf{b}(z_0) = \mathbf{P}(z_0, z_0)\mathbf{b}(z_0)$, since $\mathbf{P}(z_0, z_0)$ is a unit matrix. For an arbitrary value of z , we can, then, write

$$\mathbf{b}(z) = \mathbf{P}(z, z_0)\mathbf{b}(z_0). \quad (7.29)$$

This follows from the fact that $\mathbf{b}(z)$ and $\mathbf{P}(z, z_0)$ are both solutions of the same equation according to (7.24) and (7.25). Equation (7.29) shows that we can determine the vector $\mathbf{b}(z)$ at an arbitrary level z from its value $\mathbf{b}(z_0)$ at a reference level z_0 , by means of the matrix $\mathbf{P}(z, z_0)$. For this reason $\mathbf{P}(z, z_0)$ is called the *propagator matrix*. According to (7.28) and (7.29), if we have n levels, we can relate the displacement–stress vector for level n , $\mathbf{b}(z_n)$, to that of the zeroth level $\mathbf{b}(z_0)$:

$$\mathbf{b}(z_n) = \prod_{i=1}^n \mathbf{P}(z_i, z_{i-1})\mathbf{b}(z_0). \quad (7.30)$$

By changing the order of levels we can also write

$$\mathbf{b}(z_0) = \prod_{i=1}^n \mathbf{P}(z_i, z_{i-1})\mathbf{b}(z_n). \quad (7.31)$$

Equations (7.29)–(7.31) show the meaning of the propagator matrix $\mathbf{P}(z, z_0)$, which allows the determination of the displacement–stress vector $\mathbf{b}(z)$ at an arbitrary level from its value at a reference level, $\mathbf{b}(z_0)$. The form of equations (7.30) and (7.31) is especially useful when we have layered media with constant elastic parameters.

7.4 A layered medium with constant parameters

Wave propagation in a layered medium formed by n layers with constant parameters can be conveniently studied using the formulation in terms of the displacement–stress vector and the propagator matrix. According to equations (7.22) and (7.23), for each layer matrix A is constant. Let us start by considering the problem with only one dimension. Equation (7.24) is reduced to

$$\frac{dy(z)}{dz} = ay(z), \quad (7.32)$$

where a is a constant. A solution of this equation is

$$y(z) = e^{az}. \quad (7.33)$$

If we know the value of the function for a reference level z_0 , $y(z_0) = e^{az_0}$, we can write

$$y(z) = e^{a(z-z_0)}y(z_0). \quad (7.34)$$

This equation has the same form as (7.29), so we can define the propagator matrix, in this case with only one element, in the form

$$P(z, z_0) = e^{a(z-z_0)}. \quad (7.35)$$

For the general case when $b(z)$ is a vector of m elements and A is a square matrix of $m \times m$ constant elements, the solution of equation (7.24) is

$$b(z) = e^{A(z-z_0)}b(z_0). \quad (7.36)$$

The propagator matrix has the form

$$P(z, z_0) = e^{A(z-z_0)}. \quad (7.37)$$

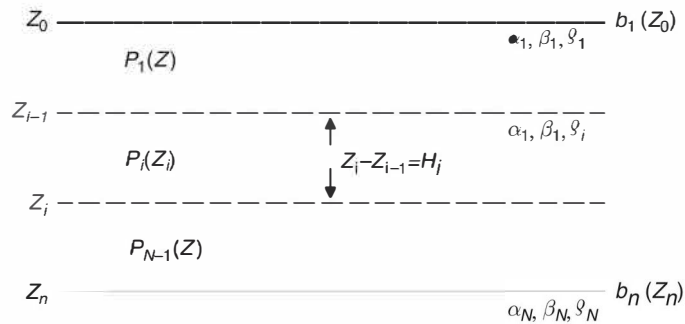


Fig. 7.3

A layered medium with constant parameters for each layer.

If we have n levels corresponding to n layers, each with constant parameters, $A_i = A(\alpha_i, \beta_i, \rho_i, c, \omega)$ (Fig. 7.3), then, according to (7.28), the propagator matrix is given by

$$P(z_n, z_0) = \prod_{i=1}^n e^{A_i(z_i - z_{i-1})}. \quad (7.38)$$

The relation between the displacement–stress vector $\mathbf{b}$ for the n th level and that for the reference (zeroth) level (7.30) for a layered medium of constant parameters is

$$\mathbf{b}(z_n) = \prod_{i=1}^n e^{A_i(z_i - z_{i-1})} \mathbf{b}(z_0). \quad (7.39)$$

Thus, the propagator matrix is given by the product of the exponential functions of the matrix A for each layer.

7.4.1 Eigenvalues and eigenvectors

The solution of equation (7.39) can be simplified by using the eigenvalues and eigenvectors analysis of matrix A (Section 4.1.1) (Lanczos, 1957). If A has eigenvalues λ_k and eigenvectors $\mathbf{u}^k$, these satisfy the equation

$$(A_{ij} - \lambda_k \delta_{ij}) u_i^k = 0. \quad (7.40)$$

The eigenvalues are solutions of the equation

$$|A_{ij} - \lambda_k \delta_{ij}| = 0. \quad (7.41)$$

The transpose matrix A^T has the same eigenvalues and eigenvectors $\mathbf{v}^k$. If we form the matrix U with eigenvectors $\mathbf{u}^k$, the matrix V with $\mathbf{v}^k$ and a diagonal matrix Λ with eigenvalues λ_k , according to the *theorem of Eckart and Young* (named after Carl Eckart and Gale Young), the matrix A and its inverse A^{-1} are given by

$$A = U \Lambda V^T, \quad (7.42)$$

$$A^{-1} = U \Lambda^{-1} V^T. \quad (7.43)$$

Equation (7.43) is called the *generalized inverse matrix*. The matrices V and U satisfy the property that the product of one by the transpose of the other is the unit matrix, $U V^T = U^T V = V V^T = V^T U = I$. Using this property and relation (7.42) and multiplying by V^T from the left-hand side in equation (7.24), we obtain

$$\frac{d}{dz} (V^T \mathbf{b}) = \Lambda (V^T \mathbf{b}). \quad (7.44)$$

Then, according to (7.36),

$$V^T \mathbf{b}(z) = e^{\Lambda(z-z_0)} V^T \mathbf{b}(z_0). \quad (7.45)$$

If we multiply by U from the left-hand side, since $UV^T = I$, we obtain

$$\mathbf{b}(z) = U e^{A(z-z_0)} V^T \mathbf{b}(z_0). \quad (7.46)$$

The propagator matrix (7.37) is now given by

$$P(z, z_0) = U e^{A(z-z_0)} V^T. \quad (7.47)$$

If we now multiply (7.46) by V^T from the left-hand side and substitute $\mathbf{w} = V^T \mathbf{b}$, we obtain an equation similar to (7.36) for the new vector $\mathbf{w}$:

$$\mathbf{w}(z) = e^{A(z-z_0)} \mathbf{w}(z_0). \quad (7.48)$$

For the vector $\mathbf{w}$ the propagator matrix is

$$Q(z, z_0) = e^{A(z-z_0)}. \quad (7.49)$$

The vector $\mathbf{w}$ is called the *wave vector* and its propagator matrix $Q(z, z_0)$ is called the *wave propagator*.

Equations (7.46) and (7.48) relate values of the vectors $\mathbf{b}$ and $\mathbf{w}$ from one level to another. They are fundamental in the solution of problems of wave propagation in stratified media. By using the generalized inverse matrix (7.43) we avoid direct calculation of the inverse of the matrix A which may be difficult in some cases. This procedure is very practical since there are standard fast methods for calculating the eigenvalues and eigenvectors of a matrix. This methodology will be used later in Chapters 16 and 18 in the resolution of the hypocentral determination and the inversion of the moment tensor.

7.4.2 The propagator matrix for SH motion

As an example, we present the determination of the propagator matrix of SH motion. According to equation (7.22), the matrix A for SH motion is given by

$$A = \begin{pmatrix} 0 & (\rho\beta^2)^{-1} \\ \rho k^2(\beta^2 - c^2) & 0 \end{pmatrix}. \quad (7.50)$$

Its transpose is

$$A^T = \begin{pmatrix} 0 & \rho k^2(\beta^2 - c^2) \\ (\rho\beta^2)^{-1} & 0 \end{pmatrix}. \quad (7.51)$$

The eigenvalues of A are found by solving the corresponding equation (7.41); we obtain $\lambda_1 = iks$ and $\lambda_2 = -iks$. For A^T the eigenvalues are the same. The matrix A is given by

$$A = \begin{pmatrix} iks & 0 \\ 0 & -iks \end{pmatrix}. \quad (7.52)$$

By solving equation (7.40) for $\mathbf{A}$ and $\mathbf{A}^T$, we obtain the following eigenvectors, corresponding to each eigenvalue:

$$\begin{aligned} \mathbf{u}^1 &= [1, \text{iksp}\beta^2], & \mathbf{u}^2 &= [1, -\text{iksp}\beta^2] \\ \mathbf{v}^1 &= [1, (\text{iksp}\beta^2)^{-1}], & \mathbf{v}^2 &= [1, -(\text{iksp}\beta^2)^{-1}]. \end{aligned}$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ are

$$\mathbf{U} = \begin{pmatrix} 1 & 1 \\ \text{iksp}\beta^2 & -\text{iksp}\beta^2 \end{pmatrix}, \quad (7.53)$$

$$\mathbf{V} = \begin{pmatrix} 1 & 1 \\ (\text{iksp}\beta^2)^{-1} & -(\text{iksp}\beta^2)^{-1} \end{pmatrix}. \quad (7.54)$$

On putting $\text{ks}\rho\beta^2 = \mathbf{a}$ and $\text{ks}(z - z_0) = d$, according to (7.47), the propagator matrix is given by

$$\mathbf{P}(z, z_0) = \begin{pmatrix} 1 & 1 \\ i\mathbf{a} & -i\mathbf{a} \end{pmatrix} \begin{pmatrix} e^{i\mathbf{a}d} & 0 \\ 0 & e^{-i\mathbf{a}d} \end{pmatrix} \begin{pmatrix} 1 & 1/(i\mathbf{a}) \\ 1 & -1/(i\mathbf{a}) \end{pmatrix}. \quad (7.55)$$

Taking the product and substituting for exponential functions sines and cosines, for real values of d , we obtain

$$\mathbf{P}(z, z_0) = 2 \begin{pmatrix} \cos d & (i/\mathbf{a}) \sin d \\ i\mathbf{a} \sin d & \cos d \end{pmatrix}. \quad (7.56)$$

In (7.56) we have assumed that d is real, or in consequence that s is real. This means that $c > \beta$, and therefore, that there are SH waves that propagate in the positive and negative directions of z , that is, waves going downward and upward. If s is imaginary, $c < \beta$, and we will have hyperbolic sines and cosines in (7.56). In this case, no SH waves are propagated in the z direction; waves are propagated in the x direction only. Furthermore, imposing the condition that their amplitudes decrease with depth, the solutions represent inhomogeneous waves, as we saw in Section 6.3.2.

7.5 SH motion in an elastic layer over a half-space

As an example, let us now apply the propagator matrix to the simple case of SH motion in a single elastic layer over a half-space (Cisternas, 1982). We use the notation $z = x_3$, $\mathbf{u}(z) = \mathbf{u}_2(z)$, and $\tau(z) = \tau_{32}(z) = \mu \, du/dz$. Subindexes 1 and 2 refer now to values of variables in the layer and the half-space, respectively (Fig. 7.4). The z coordinate is positive downward with $z = 0$ at the free surface and $z = H$ at the surface of contact between the layer and the half-space. For an arbitrary value of z , the displacement–stress vector $\mathbf{b}(z)$ (7.17) is given by

$$\begin{pmatrix} u(z) \\ \tau(z) \end{pmatrix} = \begin{pmatrix} A e^{iks z} + B e^{-iks z} \\ i\mu ks A e^{iks z} - i\mu ks B e^{-iks z} \end{pmatrix}. \quad (7.57)$$

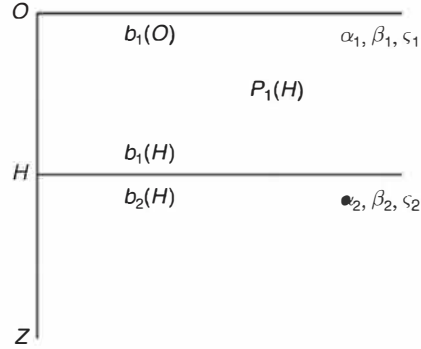


Fig. 7.4 An elastic layer over a half-space. Values of $\mathbf{b}(z)$ and $\mathbf{P}(z)$ are shown.

This vector can be expressed in terms of the product of a matrix and a vector formed by the displacement amplitudes A and B :

$$\begin{pmatrix} u(z) \\ \tau(z) \end{pmatrix} = \begin{pmatrix} e^{iks z} & e^{-iks z} \\ i\mu ks e^{iks z} & -i\mu ks e^{-iks z} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (7.58)$$

The boundary conditions are at the contact surface ($z = H$), continuity of displacement and stress (as in Section 6.3.1), and at the free surface null stress, given by

$$\tau_1(0) = 0; \quad u_1(H) = u_2(H); \quad \tau_1(H) = \tau_2(H).$$

By substitution into (7.57), we can write the displacement–stress vectors in the layer, at the free surface $\mathbf{b}_1(0)$, and at the contact surface $\mathbf{b}_1(H)$, and in the half-space at the same contact surface $\mathbf{b}_2(H)$ in the forms,

$$\mathbf{b}_1(0) = \begin{pmatrix} A_1 + B_1 \\ 0 \end{pmatrix}, \quad (7.59)$$

$$\mathbf{b}_1(H) = \begin{pmatrix} e^{id_1} & e^{-id_1} \\ ia_1 e^{id_1} & -ia_1 e^{-id_1} \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \quad (7.60)$$

$$\mathbf{b}_2(H) = \begin{pmatrix} e^{id_2} & e^{-id_2} \\ ia_2 e^{id_2} & -ia_2 e^{-id_2} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \quad (7.61)$$

where

$$\begin{aligned} a_1 &= \mu_1 s_1 k, \quad d_1 = s_1 k H \\ a_2 &= \mu_2 s_2 k, \quad d_2 = s_2 k H \end{aligned}$$

According to the definition of the propagator matrix (7.31), since $z_0 = 0$, $z_1 = H$ and, from the boundary condition on the contact surface between the layer and the half-space, $\mathbf{b}_1(H) = \mathbf{b}_2(H)$, we have

$$\mathbf{b}_1(0) = \mathbf{P}_1(H) \mathbf{b}_2(H). \quad (7.62)$$

From (7.56), the propagator matrix $P_1(H)$ can be written in the form

$$P_1(H) = 2 \begin{pmatrix} \cos d_1 & (1/a_1) \sin d_1 \\ a_1 \sin d_1 & \cos d_1 \end{pmatrix}, \quad (7.63)$$

where we have considered that s_1 is real, ($c > \beta_1$), and that SH waves propagate upward and downward inside the layer. By substitution of expressions (7.59), (7.61), and (7.63) into equation (7.62), we obtain

$$\begin{pmatrix} A_1 + B_1 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} \cos d_1 & (1/a_1) \sin d_1 \\ a_1 \sin d_1 & \cos d_1 \end{pmatrix} \begin{pmatrix} e^{id_2} & e^{-id_2} \\ ia_2 e^{id_2} & -ia_2 e^{-id_2} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}. \quad (7.64)$$

This equation relates the amplitudes in the layer, A_1 and B_1 , to those in the half-space, A_2 and B_2 . This approach will be applied to surface waves in Section 12.5.1.

7.6 The general problem

The general problem of wave propagation in a layered medium corresponds to $N - 1$ elastic layers over an elastic half-space. Each layer is characterized by its thickness H_i and elastic parameters β_i , a_i , and ρ_i . Some layers may be liquid ($\beta_i = 0$). Wave propagation in the z direction is specified by the values of s_i for SH motion and by r_i and s_i for P-SV motion (Fig. 7.5). The boundary conditions are that the stress components across the free surface ($z = 0$) are null ($\tau_{3i} = 0$). At each surface of contact, between any two layers and between the last layer and the half-space, there is continuity of the displacement and stress ($b_i(z_i) = b_{i+1}(z_i)$, $i = 1, \dots, N - 1$). In terms of the propagator matrix (7.31), the relation between the displacement-stress vector at the free surface ($z = 0$) and that at the half-space beneath the last layer ($z = z_N$) is given by

$$b_1(0) = \prod_{i=1}^{N-1} P_i(H_i) b_N(z_N), \quad (7.65)$$

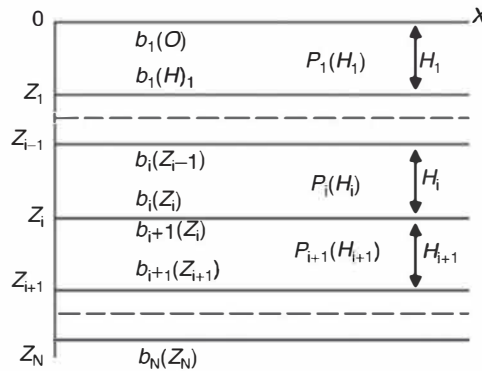


Fig. 7.5

A stratified medium with N layers of constant parameters for each layer. Values of $b(z)$ and $P(z)$ shown.

where the subindex i refers to the value at each layer from 1 to $N - 1$ and N refers to the half-space. The matrices $\mathbf{P}_i$ and vectors $\mathbf{b}_i$ for each layer depend on the parameters α , β , and ρ , frequency ω , and velocity c . According to (7.47), the propagator matrix for each layer is given by

$$\mathbf{P}_i(H_i) = \mathbf{U}_i \mathbf{e}^{\Lambda_i H_i} \mathbf{V}_i^T, \quad (7.66)$$

where Λ_i is the diagonal matrix formed by the eigenvalues of the matrix $\mathbf{A}_i$, and $\mathbf{U}_i$ and $\mathbf{V}_i$ are the matrices formed by the eigenvectors of the same matrix and its transpose.

7.6.1 SH motion

For SH motion, the matrix $\mathbf{A}$ is given by (7.50) and its transpose by (7.51). The diagonal matrix of the eigenvalues (7.52) in each layer is

$$\mathbf{A}_i = \text{diagonal}(iks_i, -iks_i). \quad (7.67)$$

According to (7.56), the propagator matrix in each layer is given by

$$\mathbf{P}_i(H_i) = 2 \begin{pmatrix} \cos(ks_i H_i) & (i/a_i) \sin(ks_i H_i) \\ i a_i \sin(ks_i H_i) & \cos(ks_i H_i) \end{pmatrix}, \quad (7.68)$$

where $a_i = ks_i \rho_i \beta_i^2$ and $k = \omega/c$. In this expression we assume that s_i is real; that is, $\beta_i < c$, and that in each layer SH waves are propagated upward and downward. If, in a particular layer, k , there is total reflection, no waves are propagated to the following layer ($k + 1$); thus, s_{k+1} is imaginary and $\beta_{k+1} > c$. If all layers have the same thickness, $H_i = H$, the problem is somewhat simplified.

7.6.2 P–SV motion

The problem of P–SV motion is more complicated since the matrix $\mathbf{A}$ has 4×4 elements (7.23). For each layer the matrix $\mathbf{A}$ is given by

$$\mathbf{A}_i = \begin{pmatrix} 0 & k^2 & (\rho_i \beta_i^2)^{-1} & 0 \\ 1 - 2\beta_i^2/\alpha_i^2 & 0 & 0 & (\rho_i \alpha_i^2)^{-1} \\ \omega \rho_i h_i & 0 & 0 & k^2(1 - 2\beta_i^2/\alpha_i^2) \\ 0 & -\rho_i \omega^2 & -1 & 0 \end{pmatrix}. \quad (7.69)$$

The eigenvalues are $\pm iks_i$ and $\pm ikr_i$. The diagonal matrix Λ is given by

$$\mathbf{A} = \text{diagonal}(ikr_i H_i, iks_i H_i, -ikr_i H_i, -iks_i H_i). \quad (7.70)$$

If r_i and s_i are real ($c > \alpha > \beta$), exponential functions of $\mathbf{A}$ according to (7.66) represent P and SV waves that propagate in the positive and negative directions of z . If r_i or s_i have imaginary values for a layer, then we have total reflection from the layer above and inhomogeneous P or SV waves that propagate in the x direction, and their amplitudes decrease with depth.

The problem is solved by forming the matrices $\mathbf{A}_i$ and their transposes $\mathbf{A}_i^T$ for each layer and determining the matrices of eigenvalues Λ_i and of eigenvectors $\mathbf{U}_i$ and $\mathbf{V}_i$. From these

matrices we find the propagator matrices $P_i(H_i)$ (7.47). The product of these matrices according to (7.65) relates the displacement–stress vector at the free surface to that of the half-space through the intermediate layers.

The matrix formulation is a powerful method for the determination of the general problem of body waves (reflected and transmitted SH and P–SV waves) in a layered medium. This is, in general, a complex problem that varies also with the depth of the focus. If the focus is at the free surface, in the half-space there are only waves traveling downward, whereas in the layers there are waves going down and up (transmitted and reflected waves). In general, the matrix formulation is very convenient for the solution of problems of wave propagation in layered media with constant parameters. In our presentation we have given only the fundamental ideas of the method; a more complete discussion can be found in Aki and Richards (1980) and Kennett (1983).

7.7 Summary

A useful model to approximate the structure of the Earth is that of a layered or stratified medium formed by parallel horizontal layers of different elastic coefficients and density. The problem of wave propagation in such a model can be solved using matrix formulation. Using Cartesian coordinates (x, z) the problem is reduced to two dimensions and is solved formulating the equation of motion in terms of the displacement–stress vector. Thus the wave equation can be formulated by a single matrix equation for the SH motion and another for the P–SV motion in which the derivatives with respect to z of the components of the displacement–stress vector are linear functions of the same components.

The solution of these equations for n layers is simplified by using the propagator matrix that allows the determination of the displacement–stress vector at an arbitrary level or layer of the medium from its value at a reference level. With this methodology, wave propagation in a layered medium with constant parameters is greatly simplified in terms of a product of exponential functions and using the eigenvalues and eigenvector analysis of the matrix with the parameters at each layer. As an example, we have applied this method to the simple case of SH motion in one layer over a half-space. Finally the more general problem for SH and P–SV motion is presented.

7.8 Problems

- P7.1 In a flat layer with constant velocity β_1 and thickness H over a half-space with constant velocity β_2 , find the ratio between the amplitudes of the SH waves at the layer and the half-space if the incident and reflected amplitudes are equal.

- P7.2 Write the P–SV matrix for a layer of constant velocity over a half-space if the thickness of the layer is 10 km, the Poisson ratio 0.25, the P wave velocity 6 km/s in the layer and 8 km/s in the half-space, the frequency of the waves $1/\pi$ Hz, the incidence angle on the layer 30° and the shear modulus $4 \times 10^{10} \text{ Nm}^2$ on the layer and $6 \times 10^{10} \text{ Nm}^2$ on the half-space.
- P7.3 For the numerical values of [Problem P7.2](#), find the value of the ratio of [Problem P7.1](#).

In previous chapters we have considered the problem of wave propagation and applied it to horizontal layered media of constant parameters that may be used as an approximation of the Earth's structure. A more realistic approximation is that of a non-homogeneous spherical medium. We will proceed towards that approximation in several steps. First of all, in non-homogeneous media the equation of motion cannot be easily reduced to the wave equation. For this reason, for the problem of the propagation of waves in inhomogeneous media or media of variable velocity, we use a different approach, called the *ray theory* approximation (Červený, 2001). As we saw in Section 5.2.1, the trajectories or paths followed by the propagation of waves are given by that of the rays or normals to the wave fronts. In many problems of seismology, we are primarily interested in these trajectories, their travel times, and the behavior of their amplitudes. It is then not necessary to find a complete solution of the wave equation; rather, it is sufficient to solve for the behavior of rays using ray theory. The early application of ray theory to the propagation of seismic rays was developed by, among others, Hans Benndorf, Zöppritz and Wiechert. Ray theory is a complex subject and here we will present its kinematic aspects in a simplified form. We will dedicate three chapters to this subject. In this chapter we begin with the fundamentals of the ray theory approximation and then apply it to the simplest case of rays propagating in planar layered media of constant velocities. In Chapter 9 we present its application to planar media of variable velocity and in Chapter 10 to spherical media. Finally, in Chapter 11 we apply these results to the determination of the structure of the Earth's interior.

8.1 The eikonal equation

In a non-homogeneous medium, as is the case in the Earth, the elastic coefficients and density are functions of the spatial coordinates. In terms of the displacements, the equation of motion (4.66) is given by

$$\frac{\partial}{\partial x_j} (C_{ijkl} u_{k,l}) + F_i = \rho \frac{\partial^2 u_i}{\partial t^2}. \quad (8.1)$$

For a liquid, for which $C_{ijkl} u_{k,l} = K u_{k,k}$, if K is a function of the coordinates, equation (8.1), for no body forces, becomes

$$\frac{\partial K}{\partial x_i} \frac{\partial u_k}{\partial x_k} + K \frac{\partial}{\partial x_i} \left(\frac{\partial u_k}{\partial x_k} \right) = \rho \frac{\partial^2 u_i}{\partial t^2}. \quad (8.2)$$

This equation cannot be reduced to the wave equation either for the cubic dilation θ or for the displacement potential ϕ , as we did in the homogeneous case (5.1 and 5.3). In terms of the cubic dilation, taking the divergence in (8.2), we obtain

$$K \frac{\partial^2 \theta}{\partial x_i \partial x_i} - \rho \frac{\partial^2 \theta}{\partial t^2} = \frac{\partial \rho}{\partial x_i} \frac{\partial^2 u_i}{\partial t^2} - \theta \frac{\partial^2 K}{\partial x_i \partial x_i} - 2 \frac{\partial K}{\partial x_i} \frac{\partial \theta}{\partial x_i}.$$

The three terms on the right-hand side are introduced by the dependence of ρ and K on the coordinates and the equation is difficult to solve. For an isotropic elastic medium, taking the derivatives in (4.68), putting $\mathbf{F} = 0$, and with λ and μ functions of the coordinates, we obtain

$$(\lambda + \mu) \frac{\partial}{\partial x_i} \frac{\partial u_k}{\partial x_k} + \mu \frac{\partial^2 u_i}{\partial x_k^2} + \frac{\partial \lambda}{\partial x_i} \frac{\partial u_k}{\partial x_k} + \frac{\partial \mu}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \rho \ddot{u}_i,$$

a rather complicated equation.

Let us now consider the problem of the propagation of P waves, that are represented by only the scalar potential ϕ , in a medium in which the velocity $v(x_i)$ varies continuously with the coordinates. If we call v_0 an initially constant value of the velocity or the value for a reference point or region, then the wave equation for such a reference region is

$$\nabla^2 \phi = \frac{1}{v_0^2} \frac{\partial^2 \phi}{\partial t^2}. \quad (8.3)$$

The solution for this equation for harmonic waves of frequency ω with constant amplitude and zero initial phase can be written just like (5.23):

$$\phi = A \exp\{i[k_0 S(x_i) - \omega t]\}, \quad (8.4)$$

where $k_0 = \omega/v_0$ is the wave number corresponding to the velocity v_0 and $S(x_i) = \text{constant}$, are the equations of the wave fronts. Let us consider the phase $\xi = k_0 S(x_i) - \omega t$, with the same value at two different times t_1 and t_2 , separated by a time interval $\Delta t = t_2 - t_1$, and wave fronts separated by a distance $\Delta S = S_2 - S_1$; then we can write

$$k_0 S_1 - \omega t_1 = k_0 S_2 - \omega t_2. \quad (8.5)$$

From these relations we can deduce that

$$v_0 = \frac{\omega}{k_0} = \frac{\Delta S}{\Delta t}. \quad (8.6)$$

In another region of the medium, where the velocity has a different value, v , the relation between two wave fronts, corresponding to the same phase, separated by a time interval Δt and distance ΔS , gives

$$v = \frac{\omega}{k} = \frac{\Delta S}{\Delta t}. \quad (8.7)$$

Let us consider now the propagation of wave fronts through a medium of smoothly continuous variable velocity $v(x_i)$. For each wave front corresponding to a constant value of the phase ζ , its time derivative must be zero:

$$\frac{d\zeta}{dt} = \frac{\partial\zeta}{\partial t} + v_i \frac{\partial\zeta}{\partial x_i} = 0, \quad (8.8)$$

and therefore,

$$\frac{\partial\zeta}{\partial t} = -v_i \frac{\partial\zeta}{\partial x_i}, \quad (8.9)$$

where the vector v_i is the velocity in the direction normal to the wave front. The gradient of ζ can be substituted by the derivative in the direction of the normal to the wave front ($\mathbf{n}$) and the vector v_i by the scalar v , the velocity in the same direction:

$$\frac{\partial\zeta}{\partial t} = -v \frac{\partial\zeta}{\partial n}. \quad (8.10)$$

According to the value of the phase $\zeta = k_0 S(x_i) - \omega t$, the gradient of ζ can be replaced by the gradient of S . Then its derivatives with respect to the normal to the wave front and with respect to time are given by

$$\frac{\partial\zeta}{\partial n} = k_0 \frac{\partial S}{\partial n}, \quad (8.11)$$

$$\frac{\partial\zeta}{\partial t} = -\omega. \quad (8.12)$$

By substituting (8.11) and (8.12) into (8.10), we obtain

$$vk_0 \frac{\partial S}{\partial n} = \omega. \quad (8.13)$$

Taking the square of (8.13) and replacing $k_0 = \omega/v_0$ and $(\partial S/\partial n)^2 = (\partial S/\partial x_1)^2 + (\partial S/\partial x_2)^2 + (\partial S/\partial x_3)^2$, we finally obtain

$$\left(\frac{\partial S}{\partial x_1}\right)^2 + \left(\frac{\partial S}{\partial x_2}\right)^2 + \left(\frac{\partial S}{\partial x_3}\right)^2 = \left(\frac{v_0}{v}\right)^2. \quad (8.14)$$

This equation is known as the *eikonal equation* (from the Greek word *eikon* for image) or the equation of *Hamilton's characteristic functions*. The equation relates the value of the square of the gradient of the wave front at a point of a medium of variable velocity to the quotient of the velocity v_0 at a reference point or the initial velocity and v at any other point. These velocities are in the direction normal to the wave fronts or along the ray's direction. The quotient $v_0/v = n$ is the refractive index (Section 6.2) at each point. Since the gradient of the wave front represents the direction of propagation or the ray's trajectory, equation (8.14) shows how this trajectory changes as wave fronts propagate through a medium of variable velocity. If we know how v varies in the medium, the eikonal equation shows us how the trajectories of rays change. Thus it provides a link between wave propagation and ray geometry.

Let us consider now the wave equation for an arbitrary point of the medium of variable velocity v , as in (8.3):

$$\nabla^2 \phi = \frac{1}{v^2} \frac{\partial^2 \phi}{\partial t^2}. \quad (8.15)$$

For a harmonic time dependence of ϕ , we can write the Helmholtz equation (5.10), in which now the wave number $k = \omega/v$ is a variable. This wave number can be expressed in terms of that of the reference region or the initial value $k_0 = \omega/v_0$, in the form $k = k_0 v_0/v$; thus, we obtain

$$\left(\nabla^2 + k_0^2 \frac{v_0^2}{v^2} \right) \phi = 0. \quad (8.16)$$

If we substitute into this equation a solution given by (8.4), we obtain

$$i \frac{\partial^2 S}{\partial x_i \partial x_i} - k_0 \frac{\partial}{\partial x_i} \frac{\partial S}{\partial x_i} + k_0 \frac{v_0^2}{v^2} = 0. \quad (8.17)$$

This equation reduces to the eikonal equation (8.14) if

$$\left| \frac{\partial^2 S}{\partial x_i \partial x_i} \right| \ll k_0 \left| \frac{\partial S}{\partial x_i} \frac{\partial S}{\partial x_i} \right|. \quad (8.18)$$

This implies that changes in the wave front are smooth. Therefore, if this condition is satisfied, for a medium of variable velocity, solutions of the eikonal equation (8.14) and, in consequence, the principles of ray theory are good approximations to the solutions of the wave equation (8.15).

8.1.1 The condition of validity

The condition of validity for the approximation of ray theory in a medium of variable velocity is given by equation (8.18). To see the significance of this condition, we start with the expression for the direction cosines of the ray (5.26):

$$\frac{\partial S}{\partial x_i} = v_i \left| \frac{\partial S}{\partial x_i} \right|. \quad (8.19)$$

If we substitute for the gradient of the wave front in (8.19) the derivatives in the direction of its normal $\mathbf{n}$ and then substitute these values into (8.18), we obtain

$$\left| \frac{\partial}{\partial x_i} \left(v_i \frac{\partial S}{\partial n} \right) \right| \ll k_0 \left| v_i \frac{\partial S}{\partial n} \right|^2, \quad (8.20)$$

where we have used

$$\frac{\partial}{\partial x_i} S = v_i \frac{\partial S}{\partial \mathbf{n}}. \quad (8.21)$$

By replacing equation (8.14) into the second term of (8.20), we obtain

$$\left| \frac{\partial}{\partial x_i} \left(v_i \frac{\partial S}{\partial \mathbf{n}} \right) \right| \ll \omega \frac{v_0}{v^2}. \quad (8.22)$$

The first term of (8.22) can be written using equation (8.13) in the form

$$\frac{\partial}{\partial x_i} \left(v_i \frac{\partial S}{\partial n} \right) = v_i \frac{\partial}{\partial x_i} \left(\frac{v_0}{v} \right) + \left(\frac{v_0}{v} \right) \frac{\partial v_i}{\partial x_i}. \quad (8.23)$$

On putting (8.23) into condition (8.22), this is satisfied if each term of the sum in (8.23) is much smaller than the right-hand term of (8.22). Taking the derivatives of v_0/v , the validity condition can be expressed by two conditions:

$$\left| \frac{\partial v}{\partial n} \right| \ll \omega, \quad (8.24)$$

$$\left| \frac{\partial v_i}{\partial x_i} \right| \ll \frac{\omega}{v}. \quad (8.25)$$

In terms of the wave length λ , these conditions can be written as

$$\frac{1}{v} \left| \frac{\partial v}{\partial n} \right| \ll \frac{1}{\lambda}, \quad (8.26)$$

$$\left| \frac{\partial v_i}{\partial x_i} \right| \ll \frac{1}{\lambda}. \quad (8.27)$$

The first condition (8.24) implies that the change in velocity in the direction normal to the wave front must be small compared with the angular frequency, or, according to (8.26), this change per unit velocity must be small compared with the inverse of the wave length. In both cases, this condition is satisfied if the frequency of the wave is large or its wave length is small. The second condition implies that the variation of the normal to the wave front (the ray's direction) must be small compared with the inverse of the wave length (8.27). This means that wave lengths must be small in comparison with the distance along which the ray changes significantly in direction. In conclusion, ray theory is a good approximation that is valid for high frequencies or short wave lengths relative to the dimensions where the velocity changes significantly in the medium.

The theory we have presented in a simplified form is known as the classical ray theory. The presence of discontinuities and strong velocity gradients in the medium, such as are found in the Earth's interior, leads to serious inadequacies and limitations of this theory. An improvement in ray theory for strong inhomogeneous media is obtained by using the WKBJ approximation (the initials are those of Gregor Wenzel, Hendrik A. Kramers, Léon Nicolas Brillouin, and Jeffreys). A presentation of this theory can be found in a more complete form in Aki and Richards (1980) and in brief form in Bullen and Bolt (1985). A more advanced treatment of the subjects of ray theory can be found in Červený (2001) and Chapman (2004).

8.2 Ray trajectories

According to the eikonal equation (8.14), the trajectories of rays or ray paths in a medium of variable velocity change depending on the variations of velocity. As wave fronts advance

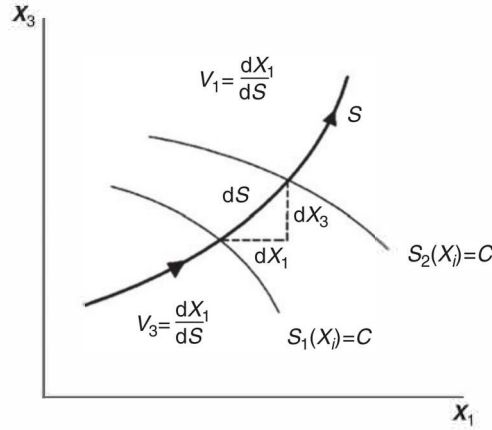


Fig. 8.1 Wave fronts S_1 and S_2 and the ray trajectory.

through the medium, they change in shape and consequently change the directions of their normals or the direction of rays. If we select a point on a wave front and, as it propagates, following the changes in the direction of its normal, we have the *trajectory of the ray* or *ray path* corresponding to that point of the wave front.

The trajectory or line of propagation of a ray is given by equation (5.27). The direction cosines of rays can be written, from (8.14) and (8.19), as

$$v_i = \frac{v}{v_0} \frac{\partial S}{\partial x_i}, \quad (8.28)$$

where v_0 is a reference velocity and $v_0/v = n$ is the refractive index at each point of the medium (Section 6.2).

If we consider the trajectory for a particular ray given by the curve s (Fig. 8.1), the direction cosines of the ray at each point can be written as

$$(v_1, v_2, v_3) = \left(\frac{dx_1}{ds}, \frac{dx_2}{ds}, \frac{dx_3}{ds} \right). \quad (8.29)$$

By substitution into (8.28), we obtain

$$\frac{\partial S}{\partial x_i} = \frac{v_0}{v} \frac{dx_i}{ds}. \quad (8.30)$$

By substituting for $\partial S/\partial x_i$ from equation (8.28) and using $d/ds = v_i \partial/\partial x_i$ we obtain an equation for the refractive index n :

$$\frac{\partial n}{\partial x_i} = \frac{d}{ds} \left(n \frac{dx_i}{ds} \right). \quad (8.31)$$

This equation can be considered to be a generalized form of Snell's law. It shows how the direction of the ray changes along its trajectory for a medium of variable velocity. In a medium of constant velocity, $v = v_0$ and $n = 1$, from equation (8.31) we obtain that the direction cosines dx_i/ds are constant along the ray's trajectory; that is, rays are straight lines.

8.3 Ray trajectories and travel times. A homogeneous half-space

Let us now consider the simplest case of ray propagation in media of constant velocity. We start by considering the propagation of rays in a homogeneous half-space. If the free surface is the horizontal plane, this is a simple way to represent a model of the Earth of planar geometry. A flat Earth is a sufficiently good approximation for small distances (less than 1000 km), for which effects of its curvature can be neglected. For the two-dimensional problem, we choose the coordinates (x, z) as defined in Chapter 7 with x along the horizontal plane and z in the vertical direction positive downward. The origin of rays is at a point (F focus or hypocenter) situated at a certain depth h from the surface and observation points are situated along the x axis on the surface. The origin of coordinates is located at the projection of the focus onto the surface, or epicenter (E) (Fig. 8.2). The depth of the focus introduces a certain spatial dimension; therefore, the application of ray theory requires that the wave lengths be much smaller than the depth h . We are interested in ray trajectories or paths and *travel times*, that is, the times that waves take to propagate from the focus to each observation point as a function of the horizontal distance $t(x)$. For a medium of constant velocity the ray trajectories are straight lines (8.31) contained in the (x, z) plane.

Times taken for rays to travel from the focus F to an observation point P for a medium of velocity v (v represents the velocity either of P or of S waves) are given (Fig. 8.2) by

$$t = \frac{(x^2 + h^2)^{1/2}}{v}, \quad (8.32)$$

Both the travel time t and the distance x can also be expressed in terms of the angle of incidence at the focus i :

$$t = \frac{h}{v \cos i}, \quad (8.33)$$

$$x = h \tan i. \quad (8.34)$$

Equation (8.32) gives the travel times $t(x)$ directly as a function of the horizontal distances, whereas equations (8.33) and (8.34) are *parametric equations* for times and distances $t(i)$

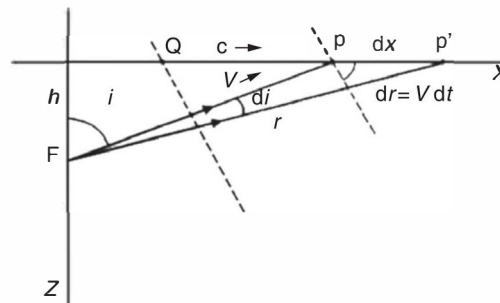


Fig. 8.2

Trajectories of two near rays in a half-space of constant velocity. The wave front advances from P to P' in time dt .

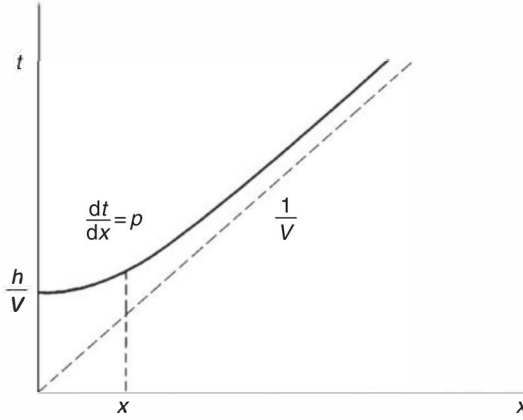


Fig. 8.3 The traveling time curve of rays in a half-space with constant velocity v for a focus at depth h .

and $x(i)$. In this case, the parameter used is the *take-off angle* i at the focus. Only for cases of simple geometry is it possible to write explicit equations for $t(x)$. However, no matter how complicated the geometry, we can write, using Snell's law, equations of parametric form. We will see that this holds also for media of variable velocity.

As we have already seen, if v is the velocity of propagation along the ray and c is the apparent velocity of the intersection of the wave front with the x axis (Fig. 8.2), then

$$c = \frac{v}{\sin i}. \quad (8.35)$$

For two rays that arrive at distances x and $x + dx$ with travel times t and $t + dt$, we can deduce easily from Fig. 8.2, where $\sin i = v dt/dx$, that

$$\frac{dt}{dx} = \frac{\sin i}{v} = \frac{1}{c} = p, \quad (8.36)$$

where p is the *ray parameter*, which, according to Snell's law, is constant along each ray. Equation (8.36) is an important relation, known as the *Benndorf relation*, between the slope of the curve of travel times at a distance x and the parameter of the ray that arrives at that distance. From (8.36), we can express $\sin i$, $\cos i$, and $\tan i$ in terms of v and p , and $\eta = 1/v$, the *slowness* or inverse of the velocity:

$$\begin{aligned} \sin i &= vp = p/\eta, \\ \cos i &= (1 - v^2 p^2)^{1/2} = v(\eta^2 - p^2)^{1/2}, \\ \tan i &= \frac{vp}{(1 - v^2 p^2)^{1/2}} = \frac{p}{(\eta^2 - p^2)^{1/2}}. \end{aligned}$$

On substituting these equations into (8.33) and (8.34), we obtain

$$t = \frac{h}{v^2(\eta^2 - p^2)^{1/2}}, \quad (8.37)$$

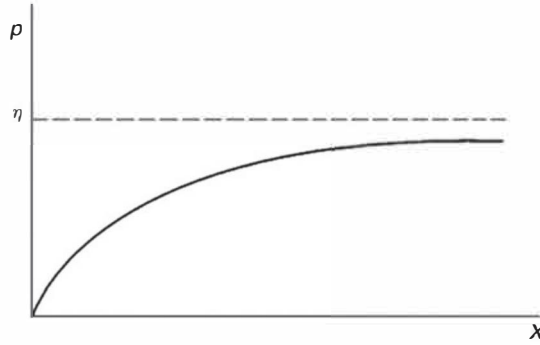


Fig. 8.4

The curve of the ray parameter with distance $p(x)$, for a half-space with constant velocity v .

$$x = \frac{ph}{(\eta^2 - p^2)^{1/2}}. \quad (8.38)$$

This is another pair of parametric equations for t and x , in terms now of the ray parameter p and the slowness η . These are very useful equations for media of variable velocity, as we will see in the [next chapter](#).

The travel time curve $t(x)$, very important in seismology, according to (8.32), is a hyperbola that tends asymptotically toward a straight line of slope $1/v$ (Fig. 8.3). If the focus is at the surface ($h = 0$), $t(x)$ is a straight line with slope $1/v$ that passes through the origin.

At every distance x there arrives a ray of a different parameter p , so the curve $p(x)$ is also very important and characteristic of each medium. For the homogeneous half-space we find from (8.38) that

$$p = \frac{\eta x}{(h^2 + x^2)^{1/2}}. \quad (8.39)$$

The curve $p(x)$ (Fig. 8.4) starts with $p = 0$ for $x = 0$, corresponding to the vertical ray upward ($i = 0$). For very large values of x the curve tends asymptotically toward $p = \eta = 1/v$. This straight line corresponds to a ray that travels horizontally ($i = 90^\circ$), which naturally exists only for $h = 0$.

Another important relation is the change in distance x with the ray parameter p . Taking the derivative with respect to p in (8.38), we obtain

$$\frac{dx}{dp} = \frac{h}{(\eta^2 - p^2)^{1/2}} + \frac{hp^2}{(\eta^2 - p^2)^{3/2}}. \quad (8.40)$$

According to this equation, dx/dp is always positive; that is, if p increases, x increases too (Fig. 8.5). For $p = 0$, dx/dp has a constant value ($h\eta$) and, for the maximum value of p ($p = \eta$), tends to infinity. For small values of p , rays are nearly vertical ($i \simeq 0^\circ$) and x varies little for changes in the angle i , whereas for values of p near $p = \eta$, rays are nearly horizontal ($i = \pi/2$) and small changes in i produce large changes in x .

Another important relation in the study of ray propagation is that of *reduced times* with the ray parameter or *tau* function $\tau(p)$, defined by

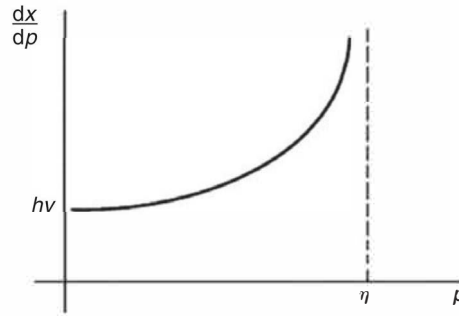


Fig. 8.5

The curve of the variation of distance with the ray parameter dx/dp , in a half-space with constant velocity u .

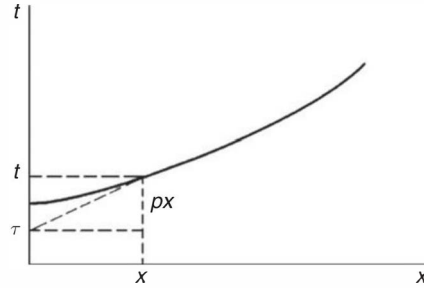


Fig. 8.6

The traveling time curve $t(x)$ and the meaning of the reduced time $\tau(p)$ (the *tau* function) for a half-space with constant velocity u .

$$\tau(p) = t(p) - px(p). \quad (8.41)$$

If we take derivatives with respect to p , we obtain

$$\frac{d\tau}{dp} = \frac{dt}{dp} - x - p \frac{dx}{dp}. \quad (8.42)$$

On substituting

$$\frac{dt}{dp} = \frac{dt}{dx} \frac{dx}{dp} = p \frac{dx}{dp},$$

we obtain

$$\frac{d\tau}{dp} = -x. \quad (8.43)$$

The slope of the curve $\tau(p)$ gives us the horizontal distance x which corresponds to the ray with parameter p . From Fig. 8.6 we can see that τ is the intersection of the tangent to the curve $t(x)$ with the t axis at a point at a distance x and with a slope p . An important property of $\tau(p)$ is that it is always a single-valued function even for cases in which $t(x)$ is a multiple-valued function.

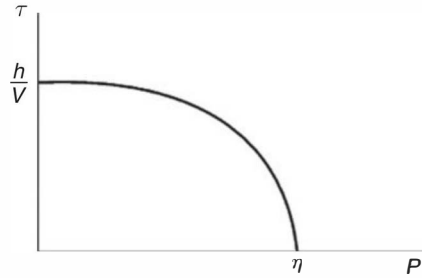


Fig. 8.7

The curve of the *tau* function (the reduced time with the ray parameter) $\tau(p)$, in a halfspace with constant velocity v .

For a homogeneous half-space, on replacing (8.37) and (8.38) in (8.41), we obtain

$$\tau(p) = h(\eta^2 - p^2)^{1/2}. \quad (8.44)$$

Figure 8.7 shows the curve $\tau(p)$. For $p = 0$, τ represents the traveling time of the vertical ray and, in the upper limit of $p(p = \eta)$, it corresponds to a horizontal ray, $\tau = 0$.

This exercise has shown the behavior of ray propagation in a homogeneous half-space and has served to define some important concepts such as the ray's trajectory or path, horizontal distance, travel time, ray parameter, apparent horizontal velocity, and reduced time. We have also seen the curves $t(x)$, $p(x)$, $\tau(p)$, and $dx/dp(p)$ that describe the characteristics of ray propagation.

8.4 A layer over a half-space with constant velocities

Another example of ray propagation in media of constant velocity is that of a layer of thickness H and velocity v' over a half-space of velocity v . This situation allows the study of ray trajectories and travel times due to the presence of a surface of contact between the layer and the half-space. The layer's thickness is a parameter that gives a spatial dimension to the model. For this reason, the application of ray theory is valid if the wave lengths are much smaller than the thickness of the layer. This example represents, in a simplified way, the situation of the Earth's crust over the upper mantle (Sections 2.6 and 11.3) for small distances, for which the flat-Earth approximation is valid.

We start with the velocity of the layer smaller than that of the half-space ($v' < v$) and the focus at the surface ($h = 0$) (Fig. 8.8). At a point P at a distance x from the focus we have three types of rays: (a) *direct rays*, traveling from F to P; (b) *reflected rays* on the surface between the layer and the half-space (FCP); and (c) *critically refracted rays* or *head waves*, that is, rays with a critical angle of incidence i_c at the contact surface that propagate a certain distance horizontally through the half-space and come back to the free surface with the same angle (FBDP) (see Section 6.2.2 and equation (6.33)). Travel times for the three types of ray are

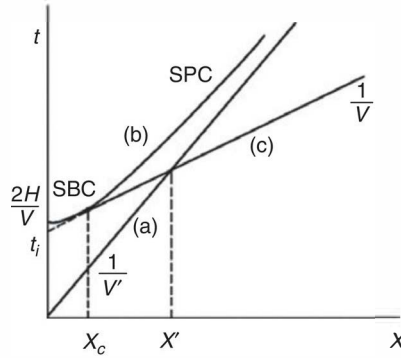


Fig. 8.9

Travel times of rays in a layer over a half-space: (a) direct rays; (b) reflected rays, SBC, subcritical and SPC, supercritical; and (c) critically refracted rays (head waves).

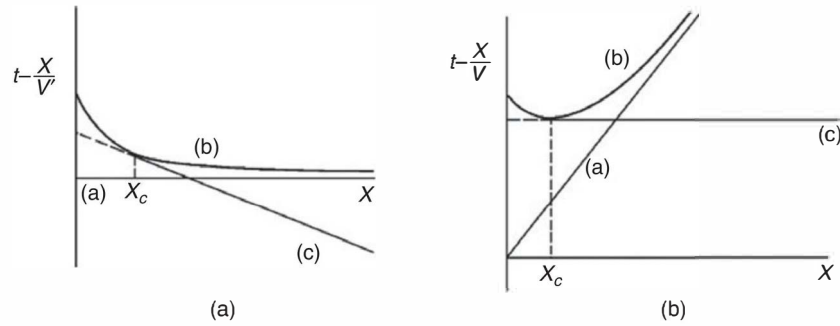


Fig. 8.10

Reduced travel time curves for a layer over a half-space: (a) the reduced velocity v' (the layer velocity), and (b) the reduced velocity v (the half-space velocity).

$$x' = 2H \left(\frac{v + v'}{v - v'} \right)^{1/2} \quad (8.50)$$

A useful representation is given by the *reduced travel times*, the ordinate representing times reduced with respect to a certain velocity v_R in the form $t - x/v_R$. Figure 8.10a shows reduced travel times with $v_R = v'$, and Figure 8.10b shows those with $v_R = v$. In both cases, this type of representation increases the differences between the slopes of the lines for the travel times of the rays. The curve corresponding to rays with velocity equal to v_R is parallel to the x axis, those corresponding to lower velocities have positive slopes, and those corresponding to higher velocities have negative slopes.

In order to study travel times of reflected rays, it is also useful to represent square times and distances (t^2, x^2). In this representation, travel times of reflected rays are straight lines with slopes $1/v'^2$ (Fig. 8.11).

For each of the three types of ray (direct, head waves, and reflected) there is a curve $p(x)$ (Fig. 8.12). For direct rays (a), it is a straight line, since the value of p is constant; $p = 1/v'$. For reflected rays (b), the curve is a hyperbola that starts at the origin and tends

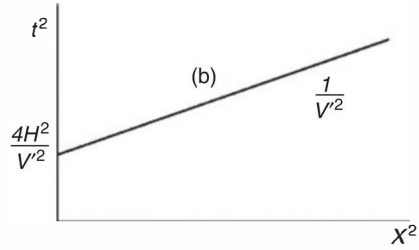


Fig. 8.11 The square of the traveling time curve (t^2, x^2) for reflected rays in a layer over a half-space.

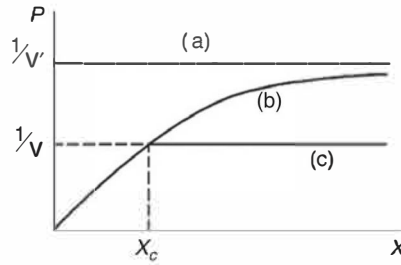


Fig. 8.12 A plot of the ray parameter versus distance $p(x)$ for a layer over a half-space: (a) direct rays, (b) reflected rays, and (c) critically refracted rays (head waves). The critical distance is x_c .

asymptotically toward the straight line of slope $1/v'$. Finally, for head waves (c), the curve is a straight line with $p = 1/v$.

The relation between travel times and the parameters of the model (H , v , and v') makes their determination easy. The layer velocity v' is the inverse of the slope of the travel times corresponding to direct rays and the velocity of the half-space v is that of critically refracted rays. Once the velocities are known, the thickness of the layer can be determined from the value of the critical distance x_c using (8.49) or from the time of intersection of head waves ($x = 0$) using (8.48). The velocity and thickness of a layer can also be determined from the slope and intercept time of reflected rays in the (x^2, t^2) representation (Fig. 8.11). This methodology, explained here only for the case of there being one layer, can be applied to the determination of velocities and thicknesses of layered media in seismic refraction profiles for the study of the Earth's crustal structure (Chapter 11). The one-layer model may be used as a first approximation to the study of the layer of sediments over a rock basement or of the crust over the upper mantle.

If the focus is not at the surface (Fig. 8.13), but rather at a depth h , equations of traveling times for direct rays (a), reflected rays (b), and critically refracted rays (c) are modified with respect to (8.45), (8.46), and (8.48) to the forms

$$t_a = \frac{(x^2 + h^2)^{1/2}}{v'}, \quad (8.51)$$

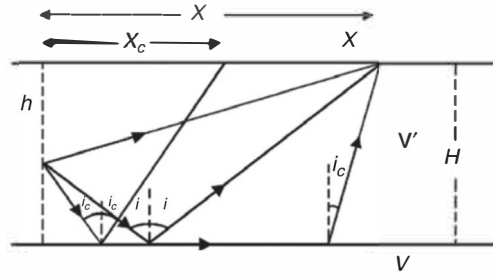


Fig. 8.13 Ray trajectories for a focus at depth h in a layer over a half-space: direct, reflected, and critically refracted rays.

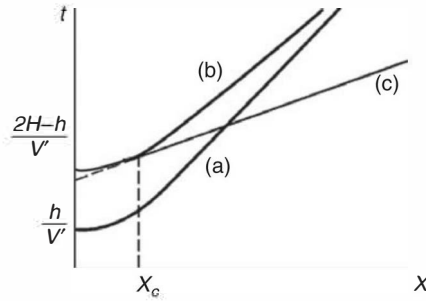


Fig. 8.14 Traveling times for a focus at depth h in a layer over a half-space: (a) the direct ray, (b) the reflected ray, and (c) the critically refracted ray (head wave).

$$t_b = \frac{[x^2 + (2H - h)^2]^{1/2}}{v'}, \quad (8.52)$$

$$t_c = \frac{x}{v} + \frac{(2H - h)(v^2 - v'^2)^{1/2}}{vv'}, \quad (8.53)$$

The critical distance is given by $x_c = (2H - h)\tan i_c$. These are more general expressions than (8.45) to (8.48), which are for the particular case of $h = 0$. The travel time curves now have the following forms (Fig. 8.14): the curve for direct rays is a hyperbola tending asymptotically toward the line of slope $1/v'$ and does not pass through the origin; the curve for critically refracted rays is a straight line of slope $1/v$, but its time intersection depends also on the focal depth; and that of reflected rays is a hyperbola as in the previous case. In these two examples we have considered that the velocity of the layer is smaller than that of the half-space. In the opposite case, there are no critically reflected rays, but only direct and reflected rays. Equations can also easily be written for the case that the focus is below the layer at the half-space ($h > H$), with layer velocity smaller or larger than that of the half-space. In the latter case only rays with incident angles less than the critical arrive at the surface.

8.5 The dipping layer

Let us consider now that the surface of the contact between the layer and the half-space is not parallel to the free surface, but has a dip angle θ . This is known as the case of a *dipping layer*. The thickness H of the layer is now measured perpendicularly from the base of the layer under the focus to the free surface. We have two different cases if we observe arrivals of rays in *down-dip* or *up-dip* manner (Fig. 8.15). Travel times of critically refracted and reflected rays are affected by the dip of the layer. For critically refracted rays and down-dip observations of a surface focus, using the results from (8.47) and that, according to Fig. 8.15a, $FC = x \cos \theta$, $EC = x \sin \theta$, $E'C = x \sin \theta \tan i_c$, and $EE' = x \sin \theta / \cos i_c$, the travel times t^- are

$$t^- = \frac{x \cos \theta}{v} - \frac{x \sin \theta \tan i_c}{v} + \frac{2H \cos i_c}{v'} + \frac{x \sin \theta}{v'}. \quad (8.54)$$

By replacing $v = v' / \sin i_c$ into the first and second terms and combining them, we obtain

$$t^- = \frac{x \sin(i_c + \theta)}{v'} + \frac{2H \cos i_c}{v'} \quad (8.55)$$

and, as a function of the velocities,

$$t^- = \frac{x \sin(i_c + \theta)}{v'} + \frac{2H(v^2 - v'^2)^{1/2}}{vv'}. \quad (8.56)$$

Because H is measured normal to the base of the layer, the depth of the layer under the focus is $H' = H / \cos \theta$.

For up-dip observations (the layer is inclined upward from the focus) (Fig. 8.15b), travel times of critically refracted rays are obtained in a similar way:

$$t^+ = \frac{x \sin(i_c - \theta)}{v'} + \frac{2H \cos i_c}{v'}. \quad (8.57)$$

If we compare equations (8.55) and (8.57) with (8.48) we find that, on putting $v' = v / \sin i_c$ in the first two, the first terms are multiplied by $\sin(i_c + \theta) / \sin i_c$ and $\sin(i_c - \theta) / \sin i_c$,

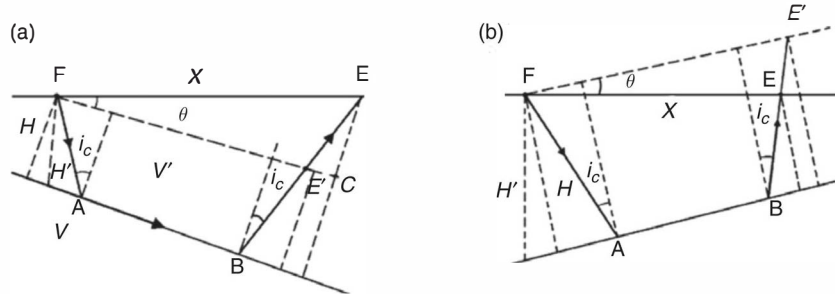


Fig. 8.15

Trajectories of critically refracted rays (head waves) for a dipping layer: (a) down-dip observations, and (b) up-dip observations.

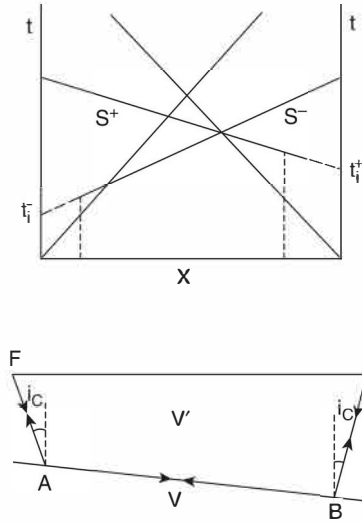


Fig. 8.16

Travel time curves in both directions (direct and inverse) for direct and critically refracted rays in a dipping layer. S^- and S^+ are the slopes of the curves of the critically refracted rays for down- and up-dip observations.

respectively. If $\theta = 0$, both factors are unity and equations (8.55) and (8.57) are equal to (8.48). These factors contain the influence of the layer's dip on the travel times.

Travel time curves for these rays are shown in Fig. 8.16, in which direct rays are also drawn. The slopes of straight lines corresponding to down- and up-dip traveling times of head waves according to (8.55) and (8.57) are

$$S^- = \frac{\sin(i_c + \theta)}{v \sin i_c}, \quad (8.58)$$

$$S^+ = \frac{\sin(i_c - \theta)}{v \sin i_c}. \quad (8.59)$$

For down-dip observations, the slope is larger than that for a flat layer of the same velocity, whereas for up-dip observations, it is smaller. If we determine the velocity of the half-space from the inverse of the traveling time curve without considering the dip of the layer, the result is larger or smaller than the real velocity. If we have observations in one direction only, we cannot detect the dip of the layer and in consequence the slope of the curve of head waves gives us only an apparent velocity for the half-space that may be smaller or larger than the real one, depending on the direction of the dip of the layer.

If we have observations in both directions, down- and up-dip, we can determine the dip of the layer and the true velocity of the half-space. The dip θ can be determined from the slopes of the critically reflected rays S^+ and S^- , using (8.58) and (8.59), and making the substitution $v' = v \sin i_c$:

$$\theta = \frac{1}{2} [\sin^{-1}(v'S^-) - \sin^{-1}(v'S^+)]. \quad (8.60)$$

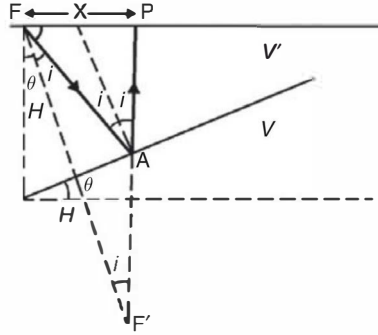


Fig. 8.17 The trajectory of a reflected ray in a dipping layer.

The velocity v' of the layer is deduced from the slope of the travel time curve of direct waves. Once the dip θ is known, the true velocity of the half-space is given by

$$v = \frac{2 \cos \theta}{S^+ + S^-}. \quad (8.61)$$

The depth along the vertical from the focus to the base of the layer can be determined from the intersection times of travel time curves of down- and up-dip head waves t_1^+ and t_1^- :

$$H' = \frac{t_1^- v'}{\cos i_c \cos \theta}, \quad (8.62)$$

$$H'^+ = \frac{t_1^+ v'}{\cos i_c \cos \theta}. \quad (8.63)$$

The determination of dips of layers and true velocities from observations of traveling times of head waves is possible only if we have observations along the same line in both directions. In the context of seismic refraction profiles these are known as *direct* and *inverse profiles* (Chapter 11).

The equation for travel times of the reflected rays in a dipping layer can easily be found by using the image method of optics ($FAP = F'AP$) (Fig. 8.17):

$$t = \frac{1}{v'} (x^2 + 4H^2 \pm 4xH \sin \theta)^{1/2}. \quad (8.64)$$

The sign depends on whether the observations are down-dip (+) or up-dip (-). The dipping layer problem can be used as a first approximation to the study of dips in the boundary between the Earth's crust and mantle or Moho (Chapter 11).

8.6 A plane-layered medium

A plane-layered medium consists of N flat parallel layers of different thicknesses H_i and different constant velocities v_i over a half-space. This is the problem we have already seen

in Chapter 7, but now we will see it from the point of view of ray propagation. For a surface focus, if the velocity increases with the depth in all layers, the traveling times of critically refracted rays at the interface at the top of layer k can be written from equation (8.48) as

$$t_k = \frac{x}{v_k} + \sum_{i=1}^k \frac{2H_i(v_k^2 - v_i^2)^{1/2}}{v_k v_i}. \quad (8.65)$$

The second term corresponds to the sum of the delay times from layer 1 to layer k . The critical distances and intersection times are given by

$$x_c^k = \sum_{i=1}^k \frac{2H_i v_i}{(v_k^2 - v_i^2)^{1/2}}, \quad (8.66)$$

$$t_1^k = \sum_{i=1}^k \frac{2H_i(v_k^2 - v_i^2)^{1/2}}{v_k v_i}. \quad (8.67)$$

Travel time curves for these rays have the form of N segments each with a slope equal to $1/v_k$ (Fig. 8.18a). The slopes have decreasing values and the critical distances and intersection times have increasing values. The reduced travel times with the reference velocity v_k are line segments with positive slopes for rays refracted at layers above layer k (smaller velocities) and negative slopes for rays refracted under it (higher velocities) (Fig. 8.18b).

If a layer has a smaller velocity than the one on top of it, this is called a *low velocity layer*. At the top of this layer there are no head waves. Head waves that arrive at the surface after traveling through the low-velocity layer are refracted at a layer with a velocity higher than that at the top of this layer. These rays are delayed by the time spent traveling through the low-velocity layers.

Travel times and distances of reflected rays at each layer interface in parametric form are given (Fig. 8.19) by

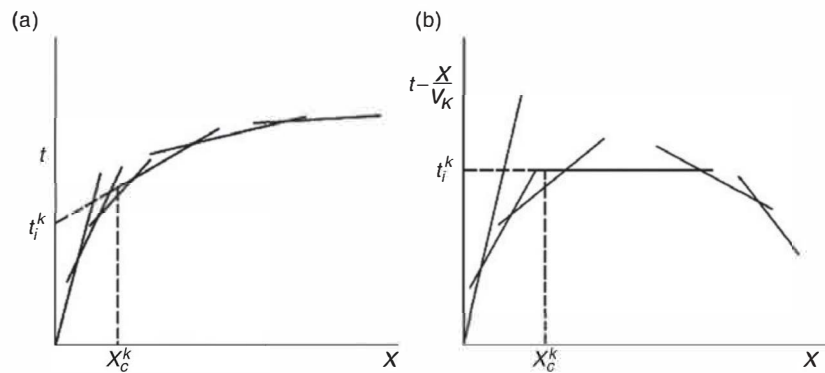


Fig. 8.18

Traveling time curves for direct and critically refracted rays (head waves) for a plane-layered medium: (a) normal traveling times, and (b) reduced traveling times (reduced velocity v_k).

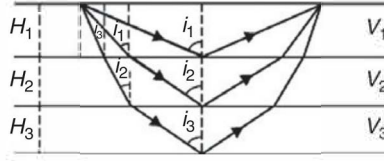


Fig. 8.19 Trajectories of reflected rays in a plane-layered medium.

$$t_k = 2 \sum_{i=1}^k \frac{H_i}{v_i \cos i_i}, \quad (8.68)$$

$$x_k = 2 \sum_{i=1}^k H_i \tan i_i. \quad (8.69)$$

According to (8.37) and (8.38), these two equations can also be written in terms of the ray parameter p_k and slowness in each layer η_i :

$$t_k = 2 \sum_{i=1}^k \frac{H_i}{v_i (\eta_i^2 - p_k^2)^{1/2}}, \quad (8.70)$$

$$x_k = 2 \sum_{i=1}^k \frac{p_k H_i}{(\eta_i^2 - p_k^2)^{1/2}}. \quad (8.71)$$

The angles of incidence of rays that pass from one layer to another, according to Snell's law, satisfy the relation

$$\frac{\sin i_i}{v_i} = \frac{\sin i_{i+1}}{v_{i+1}}. \quad (8.72)$$

Using this relation, equations (8.68) and (8.69) can be written solely in terms of the take-off angle at the focus. Therefore, for each reflected ray, we can find the traveling time and corresponding distance in terms solely of the take-off angle at the focus or of the ray parameter. Plane-layered media can be used as simplified models for the Earth's structure in the crust and upper mantle (Chapter 11).

The parametric expressions for reflected waves for a plane-layered medium (8.68)–(8.69) can also be generalized to other geometries of layers or blocks with both horizontal and vertical boundaries and also with dipping boundaries separating blocks of different constant velocities. These expressions can also be written for cases in which the focus is not at the surface. In all cases it is possible to write parametric expressions for distance and travel times in terms of the take-off angle at the focus, using Snell's law. This analysis is part of the technique called *ray tracing* in which travel times and amplitudes are calculated by following the trajectories of the rays in complex structures of constant and variable velocities (Červený, 2001). Once the take-off angle of the ray at the focus has been fixed, its trajectory is specified by Snell's law and travel times and distances can be determined. However, this is not possible for critically refracted rays or head waves, since rays that leave

the focus with the same take-off angle arrive at different distances. This approach can be used for models of the Earth's crust and upper mantle, as will be seen in [Chapter 11](#).

8.7 Summary

A realistic model of the Earth's structure should take into account its spherical shape and that its material is non-homogeneous. However, the equation of motion for media of variable elastic coefficients and density is not reducible to the wave equation. Thus we need to solve the problem of wave propagation in media with variable velocity in a different way. One way to do this is using the ray theory approximation which is valid for high frequency waves. Ray theory provides information on ray paths and travel times in media of variable velocity, which is very important in seismology.

We begin with the application of ray theory to waves in a planar medium of constant velocity where rays are straight lines. This gives us a first simple case of ray paths and travel times. A second case is that of a horizontal layer over a half-space both with constant velocity. A further complication is when the contact between the layer and the half-space is not parallel to the free surface or the dipping layer. A generalization is presented for ray paths and travel times in a medium formed by N layers over a half-space of constant velocities. These examples may serve as first approximations for models of the crust and upper mantle.

8.8 Problems

- 8.1 For a medium formed by two layers of the same thickness H (1 and 2) and a half-space (3) with velocities $V_2 < V_1 < V_3$, determine the equations for the traveling times of a surface focus for the direct and critically refracted waves. Draw the rays' trajectories and travel time curves using units of x/H and H/V_1 .
- 8.2 Consider a medium like that in [Problem 8.1](#) with velocities V , $3V$, and $2V$, with a focus at a depth $3H$. Write the parametric equations for x and t as functions of i_0 (the take-off angle at the focus) for the rays that reach the surface without reflection. Give the range of values of i_0 for each ray.
- 8.3 Given the following first arrival times x (km), t (s):

x	t	x	t	x	t
9	6	50	30	130	45
20	13	80	36	190	50
30	20	100	40	230	53

determine the model of plane layers with a constant velocity that satisfies the data.

- 8.4 Consider a medium with one layer of thickness H over a half-space, of velocity $2V$; the layer is divided vertically into two regions with velocities $V/2$ for $x < H$ and V for $x > H$. Find the distance and arrival time for the ray that crosses the vertical boundary and is reflected on the half-space with the critical angle (write expressions in terms of H and V).
- 8.5 The Earth consists of a layer of thickness 20 km and seismic wave velocity 6 km/s on top of a medium of speed of propagation 8 km/s. A seismic focus is located at a depth of 10 km. Calculate the difference in travel times between the reflected and the critical refracted waves observed on the surface at a distance of 150 km from the epicenter.
- 8.6 In a seismogram, the S – P time difference is equal to 5.31 s, and corresponds to a regional earthquake that occurred at a depth $h = 2H$, where H is the thickness of the crust. Given that the Earth is formed by a layer of constant P wave velocity of 3 km/s, that below it there is a semi-infinite medium of double that speed of propagation, and that the Poisson ratio is 0.25, determine:
- The expression for the travel time of the P and S waves.
 - The epicentral distance for an emerging P wave with a take-off angle of 30° at the focus.
- 8.7 A semi-infinite medium consists of two media of velocities V and $3V$ separated by a vertical surface. In the first medium there is a focus of seismic waves at a depth a below the free surface and at the same distance a from the surface separating the two media. Write the expressions for the direct, reflected, and transmitted waves arriving at the free surface, and plot the travel-time curve (t, x) in units of a/V and V (neglecting waves with more than a single reflection).

In the [previous chapter](#) we have seen the principles of ray theory and its application to layered media of constant velocity. This application serves as an introduction to the problem of ray propagation in media of variable velocity. The general three-dimensional case consists of media with variable velocity, with dependence on the three spatial coordinates including surfaces of different geometries which may separate discontinuous changes in velocity. In this chapter we will consider only the planar problem with variation of the velocity with depth. For relatively short distances ($\Delta < 1000$ km), which we call *regional distances*, plane geometry is a good approximation and the Earth may be considered as a half-space limited by a free surface with velocity which varies only with depth. The inverse problem of determining the velocity distribution from the travel times is presented in this chapter.

9.1 A variable velocity with depth

For the problem of a half-space with variable velocity with depth we represent the vertical direction by the z coordinate positive downward and the horizontal direction by x . Rays are contained in the (x, z) plane and problems are reduced to two dimensions ([Section 7.1](#)). The velocity v_0 and angle of incidence i_0 at the surface are taken as reference values. The direction cosines for an arbitrary point on the trajectory of a ray are given ([Fig. 9.1](#)) by

$$v_x = \frac{dx}{ds} = \sin i, \quad (9.1)$$

$$v_z = \frac{dz}{ds} = \cos i, \quad (9.2)$$

where ds is an increment of the distance along the ray. Taking the x component in equation (8.31), replacing (9.1), and applying Snell's law, we obtain

$$\frac{d}{ds} \left(\frac{v_0}{v} \sin i \right) = 0, \quad (9.3)$$

where $v(z)$ and $i(z)$ are arbitrary functions of the depth. This equation indicates that $(v_0/v) \sin i$ has a constant value along the ray. Since, for $z = 0$, this constant is $\sin i_0$,

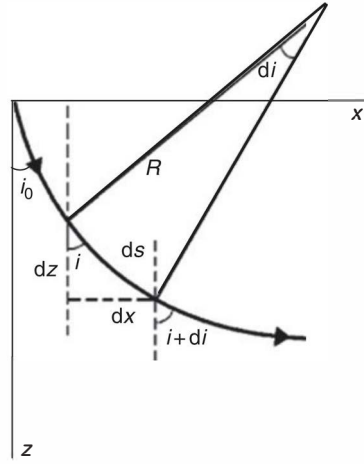


Fig. 9.1 A ray path or trajectory in a medium with variable velocity and its radius of curvature.

equation (9.3) represents Snell's law for media with velocities that are variable with depth:

$$\frac{\sin i(z)}{v(z)} = \frac{\sin i_0}{v_0} = p, \quad (9.4)$$

where, as we have already seen, p is the ray parameter, which has a constant value for each ray along its complete trajectory. If we take derivatives in equation (9.4) with respect to s (the ray's direction), then

$$\frac{d}{ds}(\sin i) = \frac{d}{ds}(pv), \quad (9.5)$$

and substituting in (9.2), we have

$$\frac{di}{ds} = p \frac{dv}{dz}. \quad (9.6)$$

This equation relates changes in the angle of incidence $i(z)$ along the ray, that is, changes in the ray trajectory, to changes in velocity.

According to (9.6), the curvature of the ray changes along its trajectory. At a point of this trajectory, the radius of curvature R_c (Fig. 9.1) is given by

$$R_c = \frac{ds}{di}. \quad (9.7)$$

Substituting equation (9.6) into this gives

$$R_c = \frac{1}{p \frac{dv}{dz}}. \quad (9.8)$$

If the velocity is constant ($dv/dz = 0$), the radius of curvature is infinite and ray trajectories are straight lines (Chapter 8). If the gradient of the velocity is constant, R is also constant

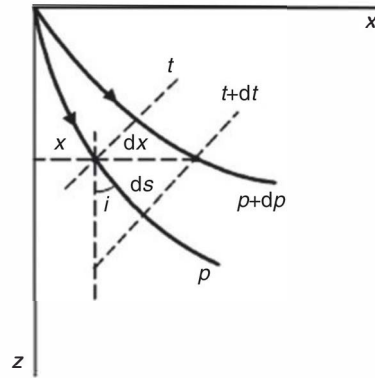


Fig. 9.2 The relation between the wave fronts and trajectories of two rays with parameters p and $p + dp$.

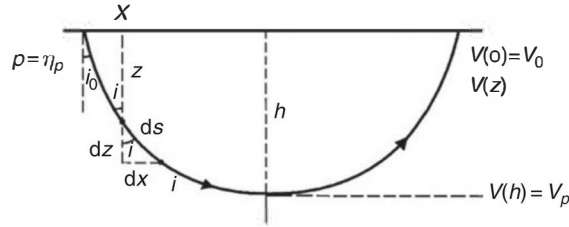


Fig. 9.3 Ray path at a medium of velocity increasing with depth starting at the surface and with turning point at depth h .

and the ray trajectory is a circle. The curvature of a ray changes along its trajectory according to the changes in the velocity gradient. The curvatures of different rays vary from each other according to their p values.

For two contiguous rays with parameters p and $p + dp$, and for two wave fronts at times t and $t + dt$, separated by a horizontal distance dx (Fig. 9.2), we can easily deduce (from $\sin i = ds/dx$ and $ds = vdt$) that

$$\frac{dt}{dx} = \frac{\sin i}{v} = p. \quad (9.9)$$

This result is similar to that obtained for a medium of constant velocity (8.36) and is known as the *Benndorf relation*.

As we have seen for the propagation of rays in media of constant velocity (Chapter 8), in the study of rays in media with variable velocities it is important to obtain expressions for the horizontal distance x , the distance traveled along the ray's trajectory s , and the travel time t . For media in which the velocity varies with depth, for a ray that starts at the surface and arrives at a certain depth z (Fig. 9.3), x , s , and t are given by the parametric equations

$$x = \int_0^z \tan i \, dz, \quad (9.10)$$

$$s = \int_0^z \frac{dz}{\cos i}, \quad (9.11)$$

$$t = \int_0^z \frac{dz}{v \cos i}. \quad (9.12)$$

If the velocity increases with depth, dv/dz is positive and the radius of curvature is also positive; that is, rays are concave upward with i increasing with depth. If rays start at the surface, i will reach $\pi/2$ at a maximum depth h , turning upward and reaching the surface, at distance x (Fig. 9.3). For rays that begin and end at the surface, the horizontal distance x , distance traveled along the ray s , and travel time t are given by twice the integrals (9.10)–(9.12), putting $z = h$. In a similar manner to what we did with equations (8.37) and (8.38), expressions (9.10)–(9.12) for rays that start and end at the surface, turning at a maximum depth h , can be written in terms of the ray parameter p and slowness η ($\eta = 1/v$), as

$$x = 2 \int_0^h \frac{p \, dz}{(\eta^2 - p^2)^{1/2}}, \quad (9.13)$$

$$s = 2 \int_0^h \frac{\eta \, dz}{(\eta^2 - p^2)^{1/2}}, \quad (9.14)$$

$$t = 2 \int_0^h \frac{\eta^2 \, dz}{(\eta^2 - p^2)^{1/2}}. \quad (9.15)$$

We can also derive an expression for reduced times or the *tau* function $\tau(p)$, by substituting (9.13) and (9.15) into (8.41):

$$\tau(p) = 2 \int_0^h (\eta^2 - p^2)^{1/2} dz. \quad (9.16)$$

We notice that expressions (9.13) and (9.15) for x and t are similar to those derived for a half-space of constant velocity (8.37) and (8.38). If the velocity increases with depth, then for each value of p there is a ray that arrives at a distance x in a time t , reaching a maximum depth h . Since, for $z = h$, $i = \pi/2$, the ray parameter $p = \eta_p = 1/v_p$, where v_p is the velocity at depth h . Along the ray, $\eta \geq p$, and there is a singularity in expressions (9.13)–(9.15) at the limit $z = h$, where $\eta = p$. According to (9.9), the slope of the travel time curve $t(x)$ equals the ray parameter p at each value of x and corresponds to the inverse of the velocity at the turning depth h ($dt/dx = 1/v_p$). This is an important relation between travel time curves and velocity distributions. Depending on the form of $v(z)$, integrals for x and t may have analytical solutions. However, only in some simple cases is it possible to find an explicit analytical function for $t(x)$. Finally, when velocity decreases with depth, dv/dz is negative, and so is the radius of curvature with rays concave downward with decreasing i with depth and limit for $i = 0$ (vertical ray).

9.2 The Lagrangian formulation

The analysis of ray propagation in media with velocities that depend solely on depth allows a simplified approach to the generalized ray theory using the Lagrangian formulation and Hamilton's canonical equations (Section 4.6). We present only a brief discussion of the problem that may serve as an introduction to more advanced treatments. The generalized formulation is very useful in the solution of problems in which the velocity varies with more than one coordinate (Červený, 2001, 104–13).

In a medium of variable velocity with depth, the travel time can be expressed, in a similar way to that in (9.12), in terms of the horizontal distance x and the differential ds along the ray (Cisternas, 1982):

$$t = \int_0^x \frac{ds}{v}. \quad (9.17)$$

The differential along the ray ds is given (Fig. 9.1) by

$$ds = (dx^2 + dz^2)^{1/2} = (1 + z'^2)^{1/2} dx, \quad (9.18)$$

where $z' = dz/dx$ and the independent variable is x , so that v is a function of z and z is a function of x . Then, equation (9.17) can be written as

$$t = \int_0^x \frac{(1 + z'^2)^{1/2}}{v(z)} dx = \int_0^x L(z, z') dx, \quad (9.19)$$

where $L(z, z')$, the time for the ray to travel the distance ds , may be considered as the Lagrangian function of the ray,

$$L(z, z') = \frac{(1 + z'^2)^{1/2}}{v(z)}. \quad (9.20)$$

Since (9.19) represents the travel time, according to Fermat's principle, this must be stationary, and, therefore,

$$\delta \int_0^x L(z, z') dx = 0. \quad (9.21)$$

As we saw in Section 4.6, this condition is satisfied if L is a solution of Lagrange's equation (4.79),

$$\frac{d}{dx} \left(\frac{\partial L}{\partial z'} \right) - \frac{\partial L}{\partial z} = 0. \quad (9.22)$$

If now we define the moment $P = \partial L / \partial z'$, then from equation (9.22) we deduce that

$$\frac{dP}{dx} = P' = \frac{\partial L}{\partial z}. \quad (9.23)$$

We can now define the Hamiltonian function $H(z, P)$ as:

$$H(z, P) = Pz' - L. \quad (9.24)$$

The differential dH is given by

$$dH = z' + dP - P' dz. \quad (9.25)$$

Since dH is a perfect differential,

$$z' = \frac{\partial H}{\partial P}, \quad (9.26)$$

$$P' = -\frac{\partial H}{\partial z}. \quad (9.27)$$

This is a system of Hamilton canonical equations of first order. Since, in our problem, the independent variable is x , we can determine dH/dx : and, using (9.26) and (9.27), we have

$$\frac{\partial H}{\partial x} = \frac{\partial H}{\partial z} z' + \frac{\partial H}{\partial P} P' = 0. \quad (9.28)$$

In consequence, H is constant with x . If we substitute the expression for the Lagrangian L from (9.20) into equation (9.24), we obtain for H

$$H = \frac{-1}{v(1 + z'^2)^{1/2}} = p, \quad (9.29)$$

where, according to (9.28), $p(x)$ is a constant. From equation (9.18) we have (Fig. 9.1) that

$$\frac{-1}{(1 + z'^2)^{1/2}} = \frac{dx}{ds} = \sin i. \quad (9.30)$$

On substituting (9.30) into (9.29), we find that p in (9.29) is the ray parameter:

$$p = \frac{\sin i}{v} = \frac{1}{c}, \quad (9.31)$$

where c is the apparent velocity of the wave front in the direction of x (Section 7.1, Fig. 7.1). Thus, the condition $H(x) = \text{constant}$ for each ray (9.28) is a formulation of Snell's law in terms of Hamilton equations. This is a different formulation of ray theory that for the simple case of a medium of variable velocity with depth has led to the same results as those we already knew using Snell's law. This type of formulation in terms of the Lagrangian and Hamilton equations is very useful for more general problems of ray theory in media of complex variable velocities.

9.3 The change of distance with the ray parameter

For a given depth distribution of the velocity, at a distance x there arrives a ray with parameter p . In order to understand the behavior of ray trajectories in media of different

velocity distributions, it is important to consider the change in distance with the change in ray parameter. Equation (9.13) with p as a variable represents distances corresponding to rays of different parameters. If we take derivatives with respect to p , we obtain

$$\frac{dx}{dp} = 2 \int_0^h \frac{dz}{(\eta^2 - p^2)^{1/2}} + 2p \frac{d}{dp} \int_0^h \frac{dz}{(\eta^2 - p^2)^{1/2}}. \quad (9.32)$$

Changing the integral over z to one over the slowness $\eta = 1/v$, putting $dz = f(\eta) d\eta$, and changing the limits of integration accordingly (for $z = 0$, to $\eta = \eta_0$; and for $z = h$, to $\eta = \eta_p$), we have

$$\frac{dx}{dp} = 2 \int_{\eta_0}^{\eta_p} \frac{f(\eta) d\eta}{(\eta^2 - p^2)^{1/2}} + 2p \frac{d}{dp} \int_{\eta_0}^{\eta_p} \frac{f(\eta) d\eta}{(\eta^2 - p^2)^{1/2}}. \quad (9.33)$$

Integrating by parts in the second term gives

$$2p \frac{d}{dp} \left[-f(\eta_0) \cosh^{-1} \left(\frac{\eta_0}{p} \right) - \int_{\eta_0}^{\eta_p} \cosh^{-1} \left(\frac{\eta}{p} \right) f'(\eta) d\eta \right]. \quad (9.34)$$

Derivation with respect to p in this expression results in

$$\frac{\eta_0 f(\eta_0)}{p(\eta_0^2 - p^2)^{1/2}} + \int_{\eta_0}^{\eta_p} \frac{\eta f'(\eta) d\eta}{p(\eta^2 - p^2)^{1/2}}. \quad (9.35)$$

Substituting (9.35) into (9.33) gives

$$\frac{dx}{dp} = \frac{2\eta_0 f(\eta_0)}{(\eta_0^2 - p^2)^{1/2}} + 2 \int_{\eta_0}^{\eta_p} \frac{[f(\eta) + \eta f'(\eta)] d\eta}{(\eta^2 - p^2)^{1/2}}. \quad (9.36)$$

Now we introduce the variable ζ , which is the gradient of the velocity per unit of velocity:

$$\zeta = \frac{1}{v} \frac{dv}{dz} = -v \frac{d\eta}{dz}. \quad (9.37)$$

Since $f(\eta) = dz/d\eta$, $\eta f'(\eta) = -1/\zeta$ and

$$f(\eta) + \eta f'(\eta) = \frac{d}{d\eta} [\eta f(\eta)] = \frac{1}{\zeta^2} \frac{d\zeta}{d\eta}. \quad (9.38)$$

By substitution into (9.36) and changing the integral over η to one over z , we finally obtain

$$\frac{dx}{dp} = \frac{-2}{\zeta_0(\eta_0^2 - p^2)^{1/2}} + 2 \int_0^h \frac{\frac{d\zeta}{dz} dz}{\zeta^2(\eta^2 - p^2)^{1/2}}, \quad (9.39)$$

where, as a function of the gradient of the velocity with depth,

$$\frac{d\zeta}{dz} = -\frac{1}{v^2} \left(\frac{dv}{dz} \right)^2 + \frac{1}{v} \frac{d^2v}{dz^2}. \quad (9.40)$$

Equation (9.39) represents the change in horizontal distance x when we pass from one ray of parameter p to another of parameter $p + dp$. This is an important relation for studying the behavior of rays in a medium with a certain distribution of velocity with depth, $v(z)$. The first term of (9.39) depends on the gradient of the velocity at the surface ζ_0 and the second depends on the gradient at each depth ζ and its change with depth, $d\zeta/dz$. If changes in the gradient of the velocity are small, the first term is dominant and dx/dp is negative. This means that rays with smaller values of p arrive at larger distances. If there are large changes in the gradient of the velocity, dx/dp may become positive and there is an inversion in the horizontal distance of travel x , that is, rays with smaller values of p arrive at shorter distances.

9.4 The velocity distribution with ζ constant

A first application of equation (9.39) corresponds to the case with ζ constant and, thus, the second term is zero:

$$\frac{dx}{dp} = \frac{-2}{\zeta_0(\eta_0^2 - p^2)^{1/2}}. \quad (9.41)$$

This case is of no direct interest in seismology, but constitutes a good exercise to help the reader understand the applications of this equation. If we assign a constant value $\zeta = \alpha$ (a small positive value), then, from equation (9.37), we obtain

$$\alpha \, dz = \frac{dv}{v}.$$

By integration of this equation with the condition that, for $z = 0$, $v = v_0$, we obtain an exponential distribution of velocity with depth:

$$v = v_0 e^{\alpha z}. \quad (9.42)$$

In this distribution, both the velocity and its gradient increase exponentially with depth, which is not a property applicable to the Earth (do not confuse α with P wave velocity). From equation (9.41), integrating with respect to p between η_0 and η_p , we obtain for x ,

$$x = \frac{1}{\alpha} \int_{\eta_0}^{\eta_p} \frac{-2dp}{(\eta_0^2 - p^2)^{1/2}} = \frac{2}{\alpha} \cos^{-1} \left(\frac{\eta_p}{\eta_0} \right). \quad (9.43)$$

Since $\eta_p = p$, the ray parameter for the distance x , on solving for $p(x)$, we have

$$p = \frac{1}{v_0} \cos \left(\frac{\alpha x}{2} \right). \quad (9.44)$$

Since $p = dt/dx$, by integrating (9.44) between 0 and x , we obtain an explicit equation for the travel time $t(x)$:

$$t = \frac{2}{\alpha v_0} \sin\left(\frac{\alpha x}{2}\right). \quad (9.45)$$

This simple exercise has shown us how to obtain travel times $t(x)$ from the expression for dx/dp for a specified distribution of velocity with depth. We will see later how we can also find certain characteristics of travel time curves, using equation (9.39), without solving the integral.

9.5 A linear increase of velocity with depth

A distribution of velocity that is very useful in seismology, being applicable in many cases to the Earth's crust and mantle (Sections 11.3–11.5), is that with a linear increase with depth,

$$v(z) = v_0 + kz, \quad (9.46)$$

where v_0 is the velocity at the surface ($z = 0$) and k is the constant velocity gradient ($dv/dz = k$). A ray that arrives at a horizontal distance x penetrates to a maximum depth h (Fig. 9.4). The ray parameter p is, then, given by

$$p = \frac{1}{v_0 + kh}. \quad (9.47)$$

According to (9.8), the radius of curvature is

$$R = \frac{v_0}{k} + h. \quad (9.48)$$

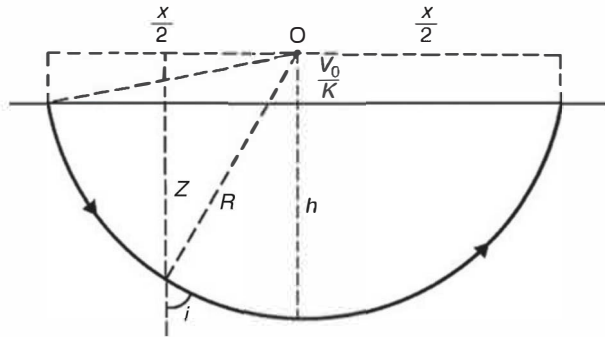


Fig. 9.4

The circular path of a ray in a medium with a linear increase of the velocity with depth.

Since R is constant, the ray's trajectory is circular with its center at a distance v_0/k above the surface (Fig. 9.4). According to the ray geometry, the radius of curvature can also be written as a function of the horizontal distance:

$$R = \left[\left(\frac{x}{2} \right)^2 + \left(\frac{v_0}{k} \right)^2 \right]^{1/2}. \quad (9.49)$$

From (9.48) and (9.49) we can deduce that the velocity at the maximum ray depth (v_m) is

$$v_m = k \left[\left(\frac{x}{2} \right)^2 + \left(\frac{v_0}{k} \right)^2 \right]^{1/2}. \quad (9.50)$$

According to Snell's law,

$$\sin i = \frac{v_0 + kz}{v_0 + kh}. \quad (9.51)$$

The travel time can be determined directly by integration; on substituting (9.51) into (9.15), we obtain

$$t = \frac{2}{k} \int_0^h \frac{k dz}{(v_0 + kz) \left[1 - \left(\frac{v_0 + kz}{v_m} \right)^2 \right]^{1/2}}. \quad (9.52)$$

And solving the integral, we obtain

$$t = \frac{2}{k} \cosh^{-1} \left(\frac{v_m}{v_0} \right). \quad (9.53)$$

Finally, using (9.50), we get the explicit equation for the travel time as a function of the horizontal distance $t(x)$:

$$t = \frac{2}{k} \sinh^{-1} \left(\frac{kx}{2v_0} \right). \quad (9.54)$$

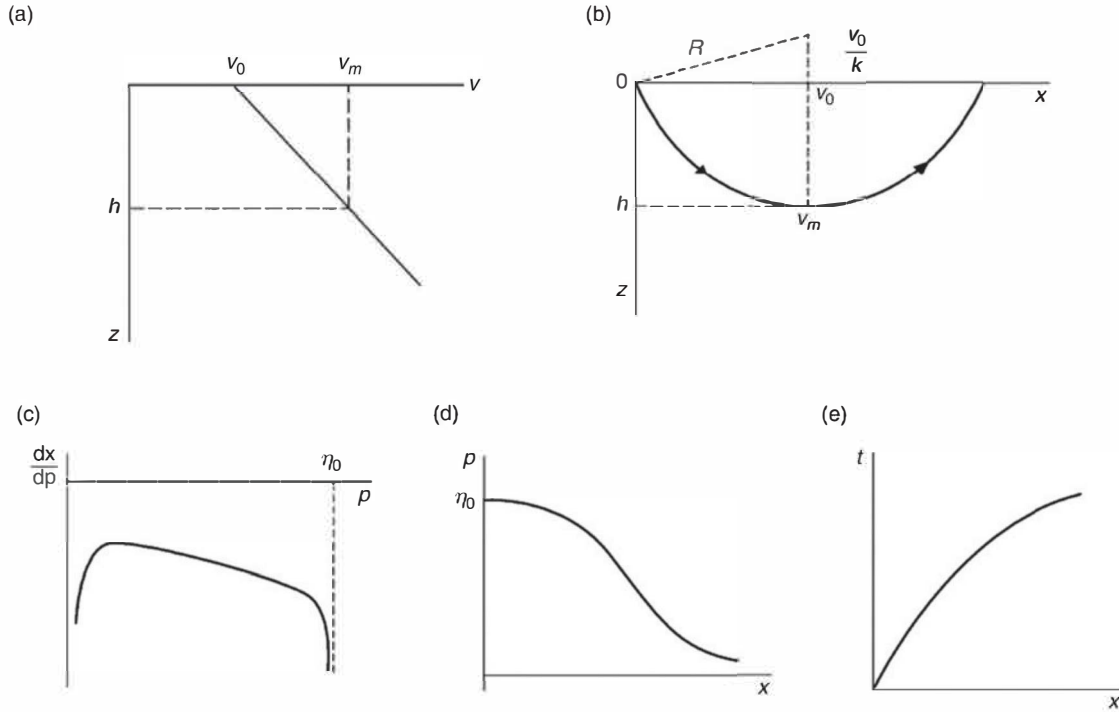
This expression can also be obtained by using equation (9.39) for dx/dp , as we did in Section 9.4. According to (9.37), for a linear increase in velocity, ζ and $d\zeta/dz$ are given by

$$\zeta = \frac{k}{v_0 + kz}, \quad (9.55)$$

$$\frac{d\zeta}{dz} = \frac{-k^2}{(v_0 + kz)^2}. \quad (9.56)$$

On substituting into (9.39), we obtain

$$\frac{dx}{dp} = \frac{-2}{kp^2(1 - v_0^2 p^2)^{1/2}}. \quad (9.57)$$

**Fig. 9.5**

A medium with a linear increase of velocity with depth: (a) the distribution of the velocity with depth; (b) a ray trajectory; (c) the variation of the distance with the ray parameter dx/dp with the ray parameter; (d) the variation of the ray parameter with distance; (e) the travel time curve.

By integration with respect to p between η_0 and η_p (the slowness at the ray's deepest point) and remembering that for each ray $\eta_p = p$, we obtain for $x(p)$,

$$x = \frac{2}{kp} (1 - v_0^2 p^2)^{1/2}. \quad (9.58)$$

Solving for p as a function of x gives

$$p = \frac{1}{\left[\left(\frac{kx}{2} \right)^2 + v_0^2 \right]^{1/2}}. \quad (9.59)$$

Since $p = dt/dx$, $t(x)$ can be obtained by integration of (9.59) between 0 and x :

$$t = \frac{1}{v_0} \int_0^x \frac{dx}{\left[\left(\frac{kx}{2v_0} \right)^2 + 1 \right]^{1/2}}. \quad (9.60)$$

The result of this integral is again expression (9.54).

Figure 9.5 shows the characteristics of a medium with a linear increase in velocity with depth, that is, one with a constant velocity gradient (Fig. 9.5(a)). Rays are circular

(Fig. 9.5(b)). The curve dx/dp (Fig. 9.5(c)) always has negative values and singularities at $p = 0$ and $p = \eta_0$. The ray parameter p decreases when x increases, with a maximum value η_0 for $x = 0$, and tends to zero for x infinite (Fig. 9.5(d)). The travel time curve $t(x)$ has a positive slope with value $dt/dx = \eta_0$, for $x = 0$, that decreases monotonically as x increases (Fig. 9.5(e)).

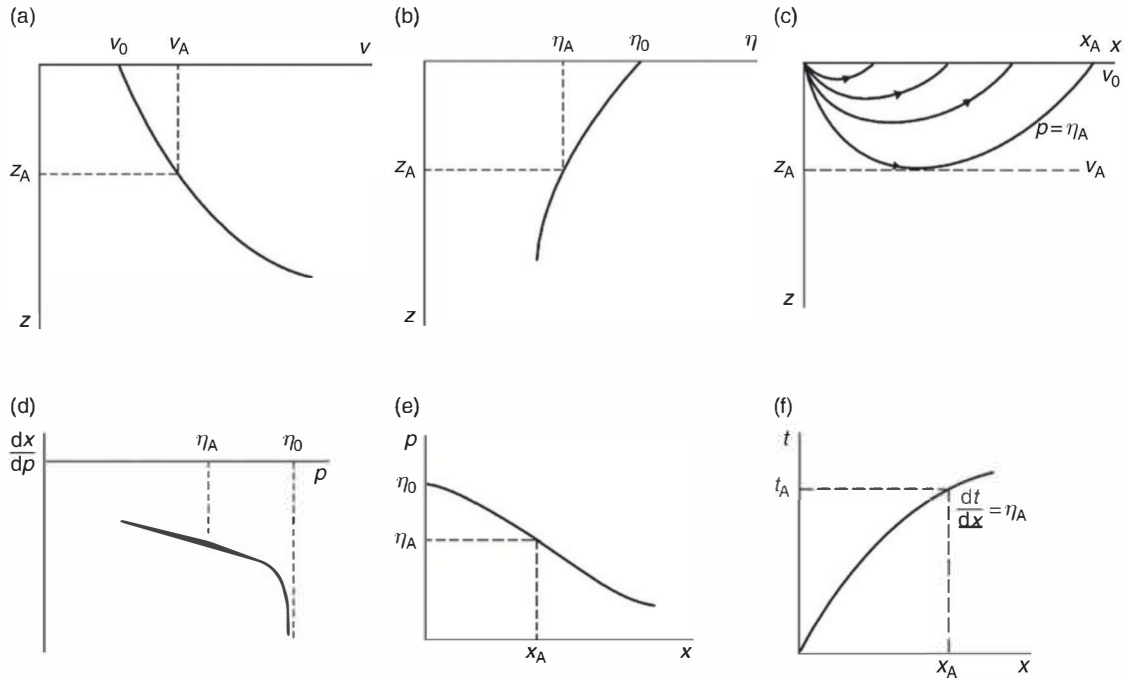
9.6 Distributions of velocity with depth

Equation (9.39) may be used to find the behavior of rays for different distributions of velocity with depth without solving completely for expressions of $p(x)$ and $t(x)$. This equation provides us with information on ray trajectories and travel times for general types of velocity distributions. The first step is to relate the characteristics of the velocity distribution $v(z)$ and its gradient dv/dz to values of ζ and $d\zeta/dz$, and thus find the general form of the curve dx/dp as a function of p . Integration of this equation gives us the curve for $p(x)$, that is, the behavior of the ray parameter with distance. Since $p = dt/dx$, a new integration gives us the travel time curve $t(x)$. Also, from knowledge of the velocity gradient, we obtain the radius of curvature (9.8) and the general characteristics of ray trajectories. In this way, even if we do not know the analytical expression for the velocity distribution with depth, knowing some of its characteristics we can find some properties of ray trajectories and travel times. We will study three cases corresponding to velocity distributions with a gradual increase, rapid increase, and decrease of velocity with depth, which are observed in the Earth's interior.

9.6.1 A gradual increase of velocity

A useful distribution of velocity with depth applicable to regions of the Earth's interior is a gradual increase with a small gradient that changes very slowly. This distribution is known in seismology as the *normal distribution of velocity*. Since dv/dz is small and positive, ζ has the same characteristics (9.37). Since there are only very small changes in velocity gradient, according to (9.40), $d\zeta/dz$ is very small. Therefore, in (9.39), the first term is the most important and dx/dp is negative. In this aspect, this case is similar to that of constant ζ (Section 9.4). The velocity distribution with constant gradient of Section 9.5, for small values of gradient, is a particular case of the normal distribution.

The most important properties of this distribution are represented in Fig. 9.6 for a surface focus. The velocity $v(z)$ increases gradually with depth (Fig. 9.6a) from a value v_0 at the surface; the gradient of velocity changes very slowly with depth. The slowness η decreases with depth from a value η_0 at the surface (Fig. 9.6b). The rays are curved, concave upward since the radii of curvature (9.8) are positive and large, dv/dz being small (Fig. 9.6c). For each distance, the ray parameter p equals the slowness at each deepest point $p = \eta_p$. A generic ray that arrives at x_A , turns at its deepest point z_A , and has a parameter $p = \eta_A$, and has a velocity at that point of $v(z_A) = v_A = 1/\eta_A$.

**Fig. 9.6**

A medium with a normal distribution of velocity with depth: (a) the distribution of velocity with depth; (b) the distribution of slowness (the inverse of velocity) with depth; (c) ray trajectories; (d) the variation of the distance with the ray parameter dx/dp ; (e) the variation of the ray parameter with distance; (f) the travel time curve.

In this distribution of velocity, dx/dp is always negative and varies with p from the singularity at $p = \eta_0$ that corresponds to $x = 0$, to a limit value for very small values of p that correspond to distances that tend to infinity (Fig. 9.6d). The limit, $p = 0$, cannot be considered since it corresponds to a vertical ray that does not arrive at the surface (or arrives at an infinite distance). For all other values of p , rays arrive at a finite distance. The integral of this curve gives us the relation between p and x (Fig. 9.6e). For distances between zero and infinity, p varies between η_0 and 0. Since $p = dt/dx$, by integration of the curve $p(x)$ we obtain the travel time curve $t(x)$ (Fig. 9.6f). At the origin the slope of $t(x)$ is $dt/dx = \eta_0$, inverse of the surface velocity v_0 . As x increases, the slope of $t(x)$ decreases monotonically. The travel time curve has, then, the form of a curve with a continuously decreasing slope, that is, d^2t/dx^2 is negative.

9.6.2 A rapid increase in velocity

Let us consider a distribution of velocity such that, at a certain depth, the velocity passes from a gradual increase to a rapid increase and, at a further depth, returns to a gradual increase (Fig. 9.7a). This situation is found in the Earth in the upper mantle and at the boundary between the outer and inner core (Chapter 11). The zone of rapid increase is located between depths z_A and z_B and there the velocity varies from v_A to v_B (the slowness

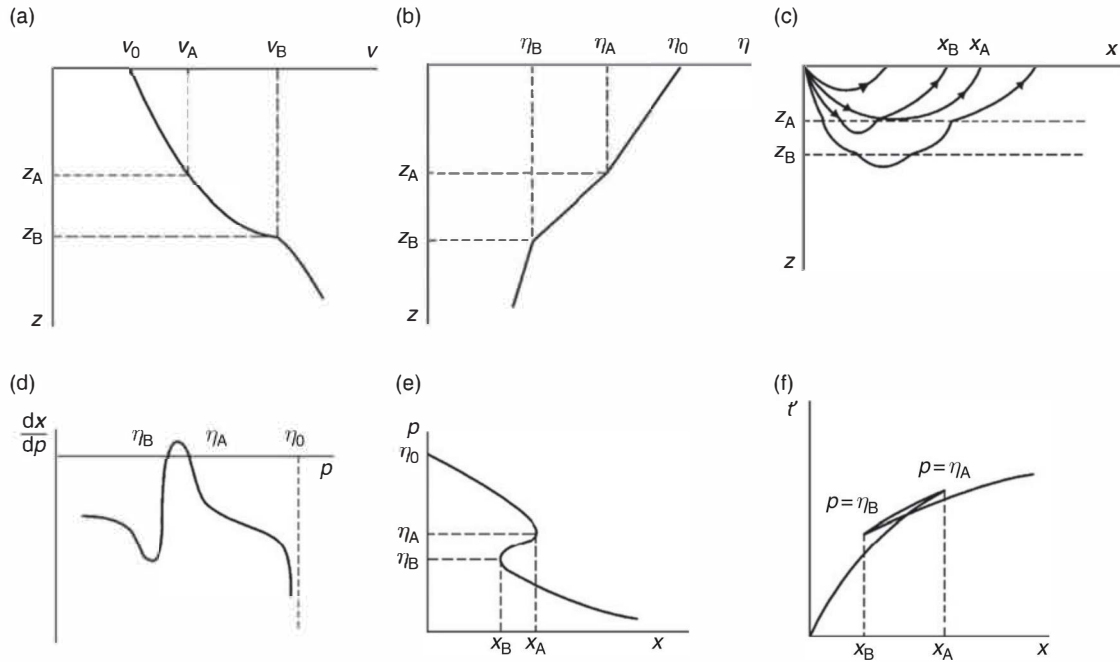


Fig. 9.7

A medium with a zone of a rapid increase of velocity with depth: (a) the distribution of velocity with depth; (b) the distribution of slowness (the inverse of velocity) with depth; (c) ray trajectories; (d) the variation of distance with the ray parameter dx/dp ; (e) the variation of the ray parameter with distance; (f) the travel time curve.

varies from η_A to η_B). The gradient in this zone is larger than those for depths above and below, so that it changes rapidly at both ends (Figs. 9.7a and b). From the origin to a distance x_A , rays do not penetrate the zone of rapid increase and have the properties of a normal distribution (Section 9.6.1).

Since the radius of curvature depends on the inverse of the velocity gradient (9.8), rays inside the zone of a rapid increase in velocity have smaller radii (Fig. 9.7c). Owing to this change, a ray that turns there may arrive at a distance x_B smaller than x_A . To understand this situation, let us consider the curve dx/dp . If, for a range of values of p near the value η_A , the change in the velocity gradient is sufficiently large that $d\zeta/dz$ is large and positive, according to (9.39), then dx/dp becomes positive (Fig. 9.7d). The curve for $p(x)$ (Fig. 9.7e) now has a range of values of p between η_A and η_B , for which x decreases when p decreases. Rays with these values of p that penetrate deeper arrive at shorter distances (Figs. 9.7c and e). For rays that turn at greater depth ($z > z_B$), where the velocity distribution is again normal, p decreases with increasing x .

The travel time curve $t(x)$ has its first part (between 0 and x_A), just like that in the form of the normal distribution (Fig. 9.7f). For distances corresponding to rays with values of p for which dx/dp is positive (between η_A and η_B), distances decrease from x_A to x_B and dt/dx increases with x (d^2t/dx^2 is positive). These rays form a second branch of the travel time curve for which the curvature has changed sign; now it is concave upward. Rays that

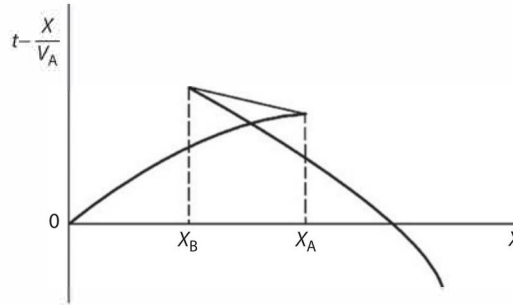


Fig. 9.8

The reduced travel time curve for a medium with a zone of a rapid increase of velocity with depth. The reduced velocity v_A corresponds to the beginning of the rapid increase.

penetrate to greater depths have negative values of dx/dp and arrive at greater distances as p decreases. On the travel time curve (Fig. 9.7f), these rays correspond to a third branch with negative curvature, concave downward, like the first. According to the $p(x)$ curve (Fig. 9.7e), for distances between x_B and x_A there exist three values of p for each distance x . This means that three different rays arrive at each distance. This results in the triplication of the travel time curves with *cusps* at their ends, which represent a focusing of rays known as a *caustic*. There is a branch of positive curvature (upward concavity), called a *retrograde branch*, since rays that penetrate deeper arrive at shorter distances ($dp/dx > 0$). The presence of this triplication of travel times is a typical feature of velocity distributions for which there is a zone of rapid increase in velocity with depth with a pronounced change in the velocity gradient. If the change in velocity gradient is not very pronounced, dp/dx does not become positive and there is no retrograde branch in the travel time curve. In this case, there is at a certain distance a change in the slope of the travel time curve which is formed by two branches, both with downward concavity.

In the reduced travel time representation, the curve has the same characteristics (Fig. 9.8). If the reduced velocity v_A corresponds to the region above the zone of rapid increase, the first branch of the curve corresponds to rays that penetrate to a depth z_A , the second for distances between x_B and x_A is the retrograde branch with positive values of dp/dx and upward concavity, and the third is a normal branch for rays that penetrate deeper than z_B . In this representation, differences among the three branches are accentuated.

We have seen that, in this case, the travel time $t(x)$ is not a single-valued function (there is a zone of the curve for which, for each value of x , there are three values of t , that is, three rays arrive at the same distance). However, for the same case, the *tau* function $\tau(p)$ (8.41) is a single-valued function (for each value of p there is one value of τ). According to (8.41), for $x = 0$, $t = 0$, $\rho = \eta_0$, and $\tau = 0$. The slope of the curve $d\tau/dp$ is always negative (8.43) and τ increases when p decreases (Fig. 9.9). The change in the slope is given by $d^2\tau/dp^2$ which, according to (8.43), is $-dx/dp$. Thus, for values of p for which dx/dp is negative, the curve is concave upward, and, when dx/dp is positive, the curve is concave downward (Fig. 9.9). In the $\tau(p)$ curve the rapid change in velocity gradient results in a change in the sign of the curvature, corresponding to the triplication of the $t(x)$ curve.

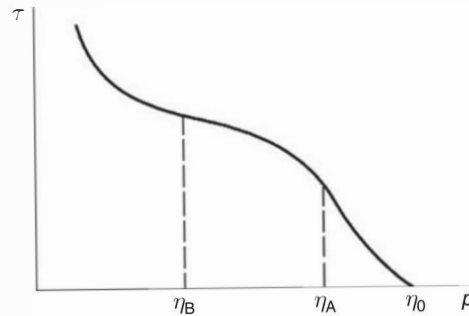


Fig. 9.9

The curve of the τ function (reduced time) with the ray parameter, for a medium with a zone of a rapid increase of velocity with depth.

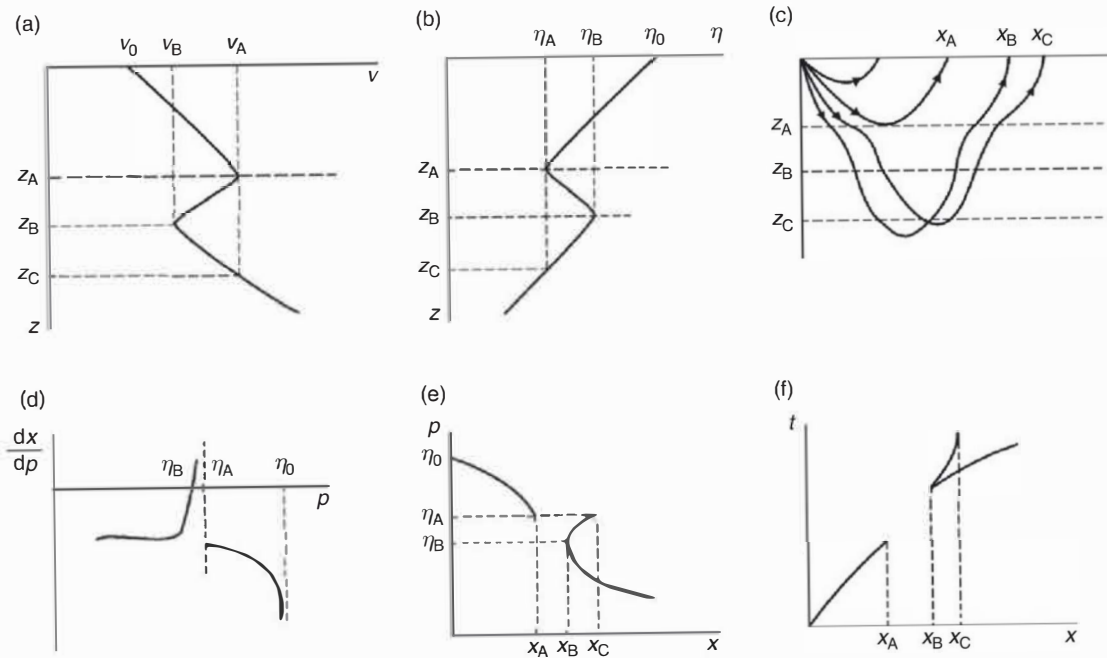


Fig. 9.10

A medium with a zone of a decrease of velocity with depth: (a) the distribution of velocity with depth; (b) the distribution of slowness (the inverse of velocity) with depth; (c) ray trajectories; (d) the variation of the distance with the ray parameter dx/dp ; (e) the variation of the ray parameter with distance; (f) the travel time curve.

9.6.3 A decrease of velocity. A low-velocity layer

We consider now a velocity distribution such that the velocity first increases with depth, then decreases at a certain depth and, further down, increases again (Fig. 9.10a). As we will see in Chapter 11, this situation exists in the Earth in the crust and upper mantle and at the

boundary between the lower mantle and the core. Between the surface and a depth z_A , the velocity increases just like it does in the normal distribution. Between z_A and z_B , the velocity decreases with depth. From z_B downward the velocity increases again. There are, then, two depths, z_A and z_C , at which the velocity has the same value v_A . The zone where the velocity is less than v_A is called a *low-velocity* layer. Rays that penetrate to depths less than z_A behave just like they do in the normal distribution, reaching a maximum distance x_A (Fig. 9.10c). Rays that penetrate into the zone where the velocity decreases with depth, since dv/dz is negative, change their curvature, which is now concave downward. Those that penetrate below z_B have again an upward concavity. According to Snell's law, rays cannot turn upwards until a depth z_C where the velocity starts to be larger than v_A .

Values of dx/dp are negative for ray parameters p between $\eta_\bullet$ and η_A , that is, for those rays that do not penetrate into the low-velocity layer. For depths between z_A and z_C , the ray parameter p is not defined, since there are no rays that turn at those depths. The curve for dx/dp is discontinuous at $p = \eta_A$ (Fig. 9.10d). Values of p less than η_A correspond to rays that turn upward at depths below z_C . Since, in the integral of (9.39), for these rays there are two zones (z_A and z_B) where the change in velocity gradient is large, $d\zeta/dz$ is also large and dx/dp may become positive, as at $p = \eta_B$ (Fig. 9.10d). In the integral of (9.39) for rays that penetrate further down, this contribution is smaller and dx/dp becomes negative again (Fig. 9.10d). In the $p(x)$ curve (Fig. 9.10e), the discontinuity of dx/dp at $p = \eta_A$ results in the existence of two rays for the same value of p , those that turn upward at z_A and arrive at x_A , and those that turn upward at z_C and arrive at a larger distance x_C . Since, for p between $\eta_B = \eta_A$, dx/dp is positive, x decreases as p decreases, corresponding to rays that arrive at distances from x_C to x_B . When dx/dp becomes negative again, x increases as p decreases (Fig. 9.10e). This change happens at a distance x_B that is always larger than x_A .

Depending on the thickness of the low-velocity zone and its magnitude, the distance interval from x_A to x_B is larger or smaller. Since, at these distances, there are no values of p , this means that no rays arrive and the interval is called a *shadow zone*. Behind the shadow zone, between x_B and x_C , there are two values of p for each value of x . Therefore, two rays arrive at each of these distances (Fig. 9.10e). For distances greater than x_C , the situation becomes normal again with just one arrival.

The travel time curve $t(x)$ (Fig. 9.10f) has, for distances up to x_A , the same form as that in the normal velocity distribution. Between the distances x_A and x_B , there is a shadow zone (δx) where no rays arrive and the curve is discontinuous. At the distance from x_B to x_C two branches correspond to the two values of p for each value of x in the $p(x)$ curve (Fig. 9.10e). The first branch starts at x_B , has the form corresponding to a normal distribution, concave downward, and continues indefinitely. The second branch starts at x_B with slope η_A , the same as that at x_A before the shadow zone, has positive curvature (concave upward) corresponding to positive values of dx/dp , and ends at x_C . This is a retrograde branch since it corresponds to the part of the $p(x)$ curve for which x decreases as p also decreases (Fig. 9.10c). A caustic appears at the cusp of the two branches. The time corresponding to the distance x_B , after the shadow zone, has a delay (δt) with respect to the prolongation of the curve from the point before it.

In conclusion, travel time curves for cases in which there are low-velocity zones exhibit a shadow zone δx , a duplication of rays, and a delay δt of arrivals after the shadow zone. The *tau* function curve $\tau(p)$ is single-valued; its slope $d\tau/dp$ increases for decreasing values of p from $\eta_\bullet$

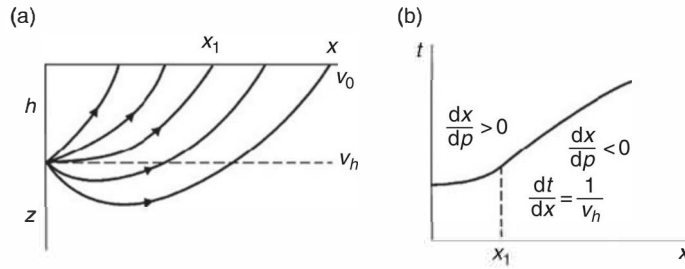


Fig. 9.11 A deep focus: (a) ray paths; (b) the travel time curve.

to η_A . At this point there is a discontinuity and, in the interval from η_A to η_B , the slope of the curve decreases with decreasing p . For values of p smaller than η_B the slope increases again.

9.7 Travel times for deep foci

Let us consider now travel times corresponding to deep foci in a medium with a normal distribution of velocity with depth. The travel time curve starts at a value of time corresponding to that of the vertical ray from the focus to the surface. For short distances, rays travel from the focus upward ($i_h > \pi/2$), at a certain distance the ray leaves the focus horizontally ($i_h = \pi/2$), and for greater distances rays leave the focus in the downward direction ($i_h < \pi/2$) (Fig. 9.11a).

The travel time curve has a first part that is concave upward corresponding to upward rays ($dx/dp > 0$). The inflection point corresponds to the distance for the ray with $i_h = \pi/2$. For larger distances, the curvature of the travel time curve is concave downward ($dx/dp < 0$) as we saw in the normal distribution (Fig. 9.6f). At the inflection point, the slope of the travel time curve corresponds to the inverse of the velocity at the focus ($dt/dx = 1/v_h$). This property can be used to find the velocity at focal depth.

9.8 Reflected rays

In media with velocity variable with depth, reflected rays are produced at the free surface and at interior surfaces where there are discontinuities in the velocity. As we did in Section 8.4 let us consider the problem of a layer of thickness H over a half-space, both with velocities which increase with depth according to a normal distribution. In the layer, velocity at the surface is v_0 and at the bottom v_1 ; velocity at the half-space begins with $v_2 > v_1$ and the focus is in the layer at depth h (Fig. 9.12a). Because the rays are curved (concave upward) the direct rays through the layer arrive at points on the free surface at distances from 0 to a maximum x_m , corresponding with the ray that turns upward at the

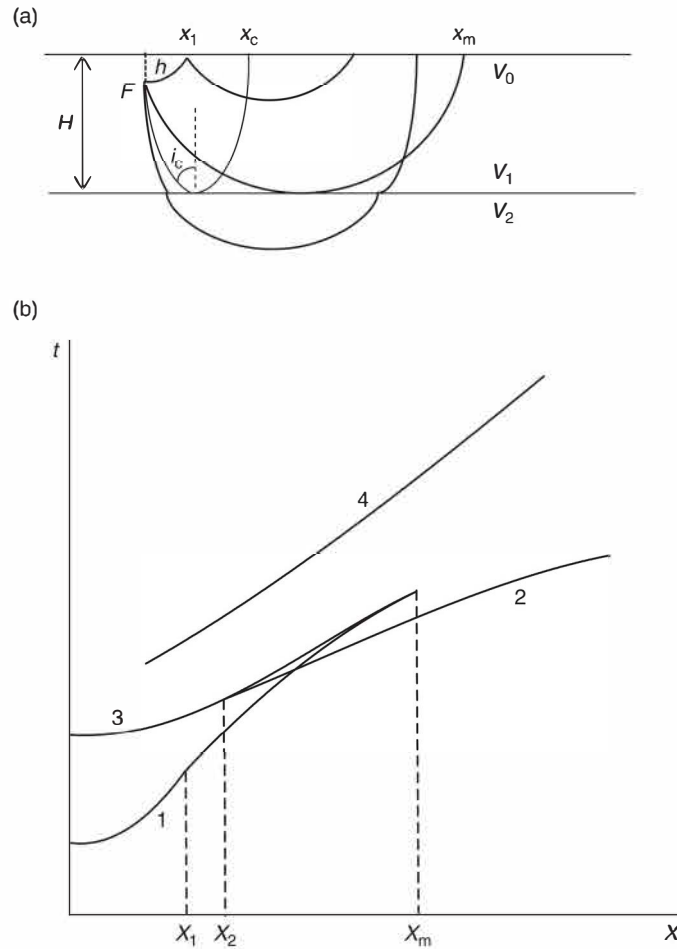


Fig. 9.12

(a) Rays in a layer over a half-space with velocities increasing with depth. (b) Travel times: (1) direct rays in the layer; (2) rays traveling through the half-space; (3) rays reflected at the boundary between the layer and the half-space; (4) rays reflected at the free surface.

bottom of the layer. In the layer there are also rays reflected at the free surface once or several times. Those reflected only once arrive at distances up to $2x_m$. Rays incident at the boundary between the layer and the half-space and reflected once arrive at the surface at distances between 0 and x_m . Rays with incident angles less than the critical angle $i_c = \sin^{-1}(v_1/v_2)$ are transmitted to the half-space, travel through that medium, and return to the free surface at distances beginning with x_c .

Figure 9.12b shows the travel time curves for the different rays. The direct ray in the layer (1) is concave upward up to distance x_1 , as in Fig. 9.11b, and concave downward for distances from x_1 to x_m . The curve for the rays traveling through the half-space (2) begins at x_c . They are the first arrivals beginning at x_2 . The curve for the reflected rays at the boundary between the layer and the half-space (3) is concave upward and extends for a distance from

0 to x_m . It has common points with curve (1) at x_m and with curve (2) at x_2 . The curve for rays reflected in the free surface (4) arrives late as they travel with lower velocities. This model can be used as an approximation for the situation in the Earth's crust and upper mantle (Sections 11.3 and 11.4).

9.9 Determination of the velocity distribution

We have seen how the form of travel time curves depends on the velocity distribution with depth. Therefore, we may be able to obtain velocity distributions from observed travel time curves. This is one of the classical inverse problems in seismology. The direct problem, that is, how to calculate theoretical travel time curves from velocity distributions, is always soluble, with greater or lesser difficulty. The inverse problem, from observed travel times to velocity distributions, involves a greater difficulty and need not always be soluble in an exact and unique way.

Let us suppose that we know the continuous travel time curve $t(x)$, in such a way that we can calculate its slope $dt/dx = p$ at each point. The distance x at which a ray of parameter p arrives, according to (9.13), changing the integral over z to one over η , is given by

$$x = 2p \int_{\eta_p}^{\eta_0} \frac{\frac{dz}{d\eta} d\eta}{(\eta^2 - p^2)^{1/2}}. \quad (9.61)$$

The limits of the integral are η_0 , the slowness for $z = 0$, and η_p that corresponds to $z = h$, the depth of the turning point of the ray that arrives at a distance x with the parameter $p = \eta_p$. Expression (9.61) represents an integral along a single ray of parameter p . Let us consider now all rays that arrive at distances from $x = 0$ to $x = x_1$, corresponding to values of p between p_0 and p_1 . For all these rays we may write the integral

$$\int_{p_1}^{p_0} \frac{dp}{(p^2 - \eta_1^2)^{1/2}},$$

where p is a variable.

If we apply this operation to both sides of (9.61), we obtain

$$\int_{p_1}^{p_0} \frac{x dp}{(p^2 - \eta_1^2)^{1/2}} = \int_{p_1}^{p_0} \int_{\eta_p}^{\eta_0} \frac{2p \frac{dz}{d\eta} d\eta dp}{[(p^2 - \eta_1^2)(\eta^2 - p^2)]^{1/2}}. \quad (9.62)$$

In this expression x and p are variables that take values between 0 and x_1 and between p_0 and p_1 , respectively. If we consider the relation between η and p , in the right-hand term of (9.62), the first integral between η_p and η_0 corresponds to a single ray that arrives at a distance x . In the second integral over p from p_1 to p_0 , η_p is a variable and the operation corresponds to integration over all rays that arrive at distances from 0 to x_1 . Considering the plane (η, p) (Fig. 9.13), the first integral from η_0 to η_p corresponds to a vertical strip. The second integral from p_1 to p_0 covers horizontally the triangle with vertices (p_0, η_0) , (p_1, η_0) , and (p_1, η_1) . Then, we can cover the same triangle by changing the order of

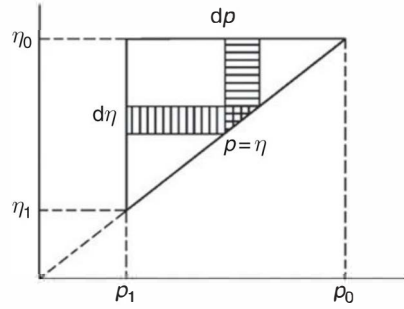


Fig. 9.13

The relation of the integrals over p and over η on the plane (p, η) .

integration, integrating first over p along a horizontal strip from p_1 to p and then vertically over η from η_1 to η_0 . The double integral is now

$$\int_{\eta_1}^{\eta_0} \int_{p_1}^p \frac{p \, dp}{[(p^2 - \eta_1^2)(\eta^2 - p^2)]^{1/2}} \frac{dz}{d\eta} d\eta. \quad (9.63)$$

The integral over p , taking into account that $p = \eta$ and $p_1 = \eta_1$ at the limits, is a known definite integral whose value is π :

$$\int_a^b \frac{x \, dx}{[(x^2 - a^2)(b^2 - x^2)]^{1/2}} = \pi.$$

The integral over η can be transformed into an integral over z from 0 to z_1 and its value is z_1 . The value of the double integral (9.63) is finally πz_1 . In the integral of the left-hand term of (9.62), changing the limits to $p_0 = \eta_0$ and $p_1 = \eta_1$ and integrating by parts, we obtain

$$\int_{\eta_1}^{\eta_0} \frac{x \, dp}{(p^2 - \eta_1^2)^{1/2}} = \left[x \cosh^{-1} \left(\frac{p}{\eta_1} \right) \right]_{\eta_1}^{\eta_0} - \int_{\eta_1}^{\eta_0} \cosh^{-1} \left(\frac{p}{\eta_1} \right) \frac{dx}{dp} dp. \quad (9.64)$$

The first term is zero, since, for $\eta = \eta_0$, $x = 0$, and, for $\eta = \eta_1$, the hyperbolic arccosine is zero. The second term can be written as an integral over x , from 0 to x_1 . The argument of the hyperbolic arccosine can be rewritten, since $p = \eta_p = 1/v_p$ and $\eta_1 = 1/v_1$, and finally we obtain

$$z_1 = \frac{1}{\pi} \int_0^{x_1} \cosh^{-1} \left(\frac{v_1}{v_p} \right) dx. \quad (9.65)$$

This equation, known as the *Herglotz–Wiechert integral* for a plane medium, was first presented by Gustav Herglotz and Wiechert in 1907. Using this expression we can calculate the velocity's distribution with depth from the travel time curve. The application of equation (9.65) consists of the following steps. The depth z_1 is the turning point (maximum depth) of the ray which arrives at a distance x_1 . At this depth the velocity is v_1 , which can be obtained from the slope of the $t(x)$ curve (at x_1 , $dt/dx = \eta_1 = 1/v_1$). To obtain the value of z_1 we have to calculate the integral of (9.65). This can be done numerically, since we can determine v_p from the slope of $t(x)$ for each value of x between 0 and x_1 (Fig. 9.14a). In this

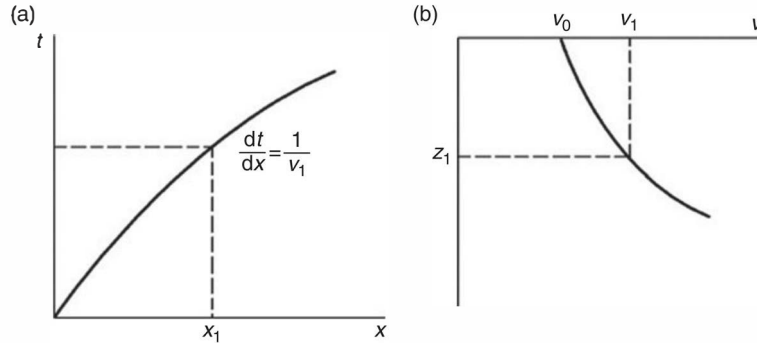


Fig. 9.14

Application of the Herglotz–Wiechert formula to obtain the velocity distribution from the travel time curve: (a) the travel time curve; (b) the distribution of velocity with depth.

way we obtain v_1 and z_1 , that is, one point of the curve $v(z)$ (Fig. 9.14b). We obtain the other points by changing the value of x_1 .

This method is valid if we know $t(x)$ and dt/dx for each value of x , and dt/dx is a monotonically decreasing function of x . This limits the method to normal velocity distributions. The method cannot be applied directly if there are velocity discontinuities, zones of very rapid increase, or low-velocity layers. However, there are modifications of the method that allow its application to such distributions. Another limitation is the need to know the travel time curve in a continuous fashion, which is not possible for observed curves that are formed by discrete values of x and t . This limitation is overcome by interpolation between observed values. For plane geometry, this method can be applied to find velocity distributions for regions of the crust or upper mantle where the above conditions are assumed to be satisfied (Chapter 11).

9.10 The energy propagated by ray beams. Geometrical spreading

From the point of view of ray theory, energy is propagated along a bundle or beam of nearby rays. Since rays may converge or diverge during their propagation due to the velocity distribution of the medium, this fact must be taken into account in order to determine the energy that arrives at a certain distance. Let us consider a point source of seismic waves from which an energy E is emitted homogeneously in all directions per unit time. The energy emitted per unit solid angle is $P = dE/d\Omega$ and the total energy per unit time is $E = 4\pi P$. For a focus at depth h , an element of solid angle $d\Omega$ encompassing all azimuths sustained by an element of take-off angle di_h corresponding to a ray beam with take-off angles between i_h and $i_h + di_h$ is given by

$$d\Omega = 2\pi \sin i_h di_h. \quad (9.66)$$

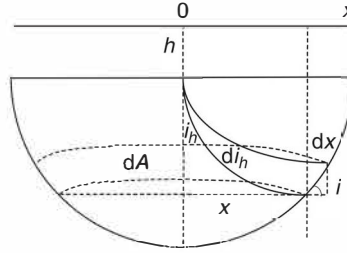


Fig. 9.15

The circular area covered by a wave front of a ray beam that leaves the focus with take-off angle i_h for a distance x .

At a horizontal distance x , the surface element of the wave front dA swept by a ray beam of solid angle $d\Omega$ (Fig. 9.15) is

$$dA = 2\pi x \cos i \, dx, \quad (9.67)$$

where i is the angle of incidence of the ray that arrives at a distance x . The energy intensity or energy per unit of wave front area is $I = dE/dA$ or, in terms of P , $I = P \, d\Omega/dA$. On substituting (9.66) and (9.67) into this, we obtain

$$I = \frac{P \sin i_h \, di_h}{x \cos i \, dx}. \quad (9.68)$$

Because the distance x reached by a ray depends on the take-off angle at the focus i_h , the distance interval can be expressed by the interval in angle of incidence $dx = (dx/di_h) \, di_h$ and, by substitution into (9.68),

$$I = \frac{P \sin i_h \, di_h}{x \cos i_h \frac{dx}{di_h}}. \quad (9.69)$$

If we differentiate with respect to x in (8.36), we obtain

$$\frac{d}{dx} \left(\frac{\sin i_h}{v_h} \right) = \frac{dp}{dx} = \frac{d^2 t}{dx^2}. \quad (9.70)$$

From (9.70) we deduce that

$$\frac{dx}{di_h} = \frac{\cos i_h}{v_h \frac{d^2 t}{dx^2}}, \quad (9.71)$$

where v_h is the velocity at the focus. By substituting in (9.69), we can express it in terms of the second derivative of the traveling time:

$$I = \frac{P v_h \tan i_h \, d^2 t}{x \cos i \, dx^2}. \quad (9.72)$$

This is an important relation between the energy intensity I at a distance x with the second derivative of the travel time $d^2 t/dx^2$. If rays arrive at the surface with angle of incidence i_0 ,

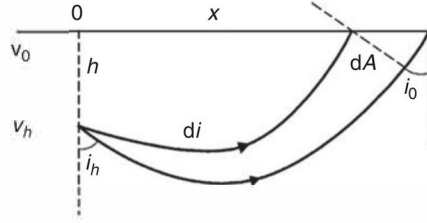


Fig. 9.16

The arrival at the surface of a wave front corresponding to a ray beam that leaves the focus with take-off angle i_h and arrives at the surface with angle of incidence i_0 .

then, by substitution into (9.68) and (9.72), we obtain the energy per unit time and per unit of wave front area that arrives at a distance x (Fig. 9.16):

$$I = \frac{P \sin i_h}{x \cos i_0 \frac{dx}{di_h}} = \frac{P v_h \tan i_h}{x \cos i_0} \frac{d^2 t}{dx^2}. \quad (9.73)$$

If the focus is at the surface, $i_h = i_0$ and $v_h = v_0$,

$$I = \frac{P v_0 \tan i_0}{x \cos i_0} \frac{d^2 t}{dx^2}. \quad (9.74)$$

In expressions (9.73) and (9.74), we can see that the energy intensity decreases with the inverse of the distance. This dependence on the factor $1/x$ is called *geometric spreading* and is due to the increase in the wave front area comprised by a ray beam as rays propagate from the focus (similar effect to what we saw in Section 5.9 for spherical waves). According to the first term of (9.73), the energy intensity depends inversely on dx/di_h . When dx/di_h is small this means that x varies little with i_h and therefore the ray density at an interval of distance dx is large. If dx/di_h is large, x varies greatly with i_h and the ray density for an interval of distance is small. According to the second term of (9.73) or (9.74), the variation in energy density with distance is related to the second derivative of the travel time. If the travel time has a uniform gradient, then the distribution of the energy density with distance is also uniform. When there are large changes in the gradient, for example at the cusps, we find an accumulation of energy at the corresponding distances. This may occur when there is a zone of rapid increase in velocity with depth, or also when there is a low-velocity zone (Figs. 9.7f and 9.10f).

9.11 Summary

We have applied ray theory to the two-dimensional problem of a planar medium with variable velocity with depth and derived equations for horizontal distance, distance along the ray, and travel times (9.13–9.15). An important relation, which gives the change of distance with the ray parameter in terms of the gradient of the velocity with depth (9.39), permits us to derive the behavior of rays and the travel times in a medium with a certain distribution of velocity with depth. As examples we have presented the linear increase of

velocity with depth, which is a particular case of the normal distribution in which there is a gradual increase of velocity with depth and media in which there is a zone with a rapid increase in the gradient and a decrease of velocity with depth. Ray paths and travel times for these cases are representative of situations in the Earth interior.

The inverse problem of the determination of the velocity distribution with depth using travel times is one of the classic inverse problems in seismology. For the planar case we have considered here, this is done using the planar formulation of the Herglotz–Wiechert integral (9.65). Applied to the Earth this allows determination of the structure of its interior in the crust and upper mantle, where the planar approximation is applicable, from the observed travel times of waves produced by earthquakes (Chapter 11). Energy propagation by ray bundles has shown the decrease of energy intensity with the inverse of distance effect known as geometric spreading.

9.12 Problems

- P9.1 In a flat medium, the velocity increases linearly with depth according to the expression $v = v_0 + kz$, where v_0 is the velocity at the surface, k is a constant, and z the depth. For a focus at depth h , calculate an expression for the take-off angle at the focus i_h in terms of the epicentral distance x , and v_0 , h , and k .
- P9.2 For Problem P9.1, if $v_0 = 4$ km/s, $k = 0.1$ s⁻¹, for a station at 171.65 km of epicentral distance, the maximum velocity of a ray that arrives at this station is 10 km/s. Estimate the focal depth and the incidence angle at the focus and at the station.
- P9.3 A layer of thickness H has a velocity distribution $v = v_0 \exp(\alpha z)$, where $\alpha < 1$. Beneath it there is a semi-infinite medium of speed of propagation $v_1 = 2v_0 \exp(\alpha H)$.
- Determine in terms of v_0 , v_1 , α , and H :
 - The distance and critical angle.
 - The time of intersection of the reflected wave.
 - The maximum distance of the direct wave.
 - Calculate the values of these parameters if $v = 1$ km/s, $H = 10$ km, and $\alpha = 0.1$ km⁻¹.
- P9.4 Beneath a layer of thickness H of velocity distribution $v = v_0 + kz$ there is a semi-infinite medium of speed of propagation $v_1 = 2(v_0 + kH)$.
- Determine expressions (as functions of the above parameters) for the critical distance, the time of intersection of the reflected wave, and the maximum distance of the direct wave.
 - For $H = 10$ km, $v_0 = 1$ km/s, and $k = 0.1$ s⁻¹, calculate these parameters and plot the travel time curves.
- P9.5 A medium is formed by a layer of thickness H and velocity v_0 over a half-space of velocity $v = v_0 + k(z - H)$.
- Determine the expressions for x and t as functions of i_0 , the take-off angle at the focus, for a focus at the surface.

- (b) Determine the relation between the angle of incidence at the focus i_0 and the depth h of penetration into the half-space.

P9.6 The travel times of P waves in a medium whose velocity is variable with depth are given by the following values; x (km), t (s):

x	t	x	t	x	t
0	0.0	200	28.8	400	41.9
50	9.6	250	33.0	450	44.2
100	17.6	300	36.4	500	46.2
150	23.9	350	39.3	550	48.1

- (a) Draw the traveling times curve.
 (b) Find the distribution of the velocity with depth.

In the study of seismic waves in the Earth for large distances ($x > 1000$ km), the planar approximation is no longer valid and we must consider the spherical shape of the Earth. The study of rays propagating inside a sphere requires certain modifications to the notions applied to the planar case in Chapters 8 and 9. We will study first the problem of constant velocity and afterwards that of velocity distributions varying with the radius. Similar situations are presented for a spherical medium to those for the planar case of Chapter 9.

10.1 The geometry of ray trajectories and displacements

Let us consider a sphere of radius r_0 and velocity depending on the radius $v(r)$. Rays propagate from a focus F on the surface to a point P that is also on the surface (Fig. 10.1). If the velocity depends solely on the radius, rays are contained on a plane that includes the center of the sphere. If the velocity increases toward the center, ray trajectories are concave toward the surface and they reach a point on the surface. Distances between two points on the surface (from F to P in Fig. 10.1) are represented by the angle Δ at the center of the sphere (the distance on the surface is $r_0\Delta$). The angle of incidence, i , is measured from the ray trajectory to the radius. The radial distance from the center to the point where a ray turns upward is r_p , corresponding to the velocity v_p (Fig. 10.1).

If the spherical medium is formed by concentric spherical layers of constant velocity (Fig. 10.2), then according to Snell's law (6.1), a ray that passes from a layer of velocity v_1 to another of velocity v_2 must satisfy

$$\frac{\sin i_1}{v_1} = \frac{\sin f}{v_2}. \quad (10.1)$$

We want to relate the angles i_1 at the base of layer 1 to i_2 at the base of layer 2. Considering the triangles POR and QOR, we can write the relation

$$r_2 \sin i_2 = r_1 \sin f. \quad (10.2)$$

By substitution in (10.1) we obtain a new formulation of Snell's law,

$$\frac{r_1 \sin i_1}{v_1} = \frac{r_2 \sin i_2}{v_2} = p, \quad (10.3)$$

where p is the ray parameter and the variables r , i , and v are in each case referred to the same point on the ray's trajectory. The units of the ray parameter p are now those of time

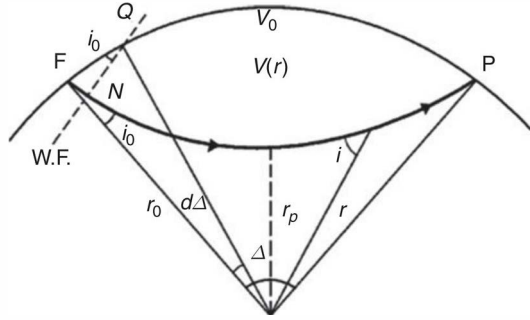


Fig. 10.1 A ray path in a sphere where the velocity increases with depth. The wave front (W.F.) advances from F to N in time dt .

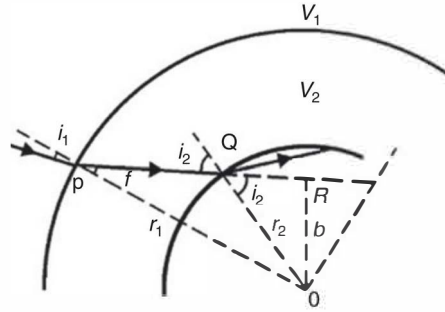


Fig. 10.2 The trajectory of a ray in three concentric spherical regions of constant velocity $u_1 < u_2 < u_3$ (Snell's law).

(seconds), whereas in the plane case they were of inverse velocity (m^{-1}s). In a spherical medium, where the velocity varies continuously with the radius, we can write for each point on a ray,

$$\frac{r \sin i}{v} = p. \quad (10.4)$$

For a wave front that advances from the focus F, within a time interval dt , a distance along the ray $FN = v_0 dt$, since $FQ = r_0 d\Delta$, we find (Fig. 10.1),

$$\sin i_0 = \frac{FN}{FQ} = \frac{v_0 dt}{r_0 d\Delta}. \quad (10.5)$$

From (10.4) and (10.5) we obtain the Benndorf relation for rays in spherical media,

$$\frac{dt}{d\Delta} = \frac{r \sin i}{v} = p. \quad (10.6)$$

At the point of the ray trajectory where it turns upward, $i = \pi/2$, corresponding to the maximum penetration of the ray and, from (10.4), $r_p/v_p = p$, where r_p and v_p are the radial distance and velocity at that point. The slowness η is now defined as $\eta = r/v$ and, in consequence, $p = \eta_p$.

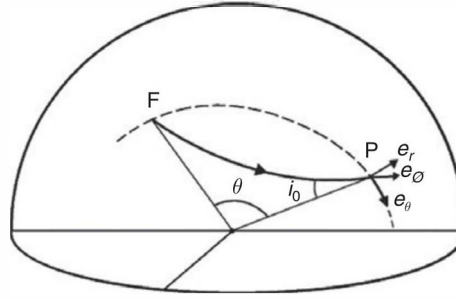


Fig. 10.3 A ray trajectory in a sphere and displacement components.

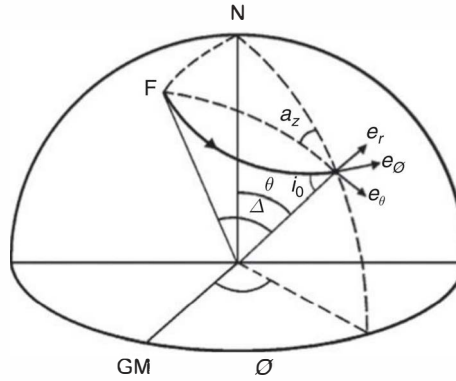


Fig. 10.4 A ray path on a sphere and displacement components at its surface referred to geocentric geographic coordinates r , θ , and ϕ (N = North Pole, GM = Greenwich Meridian).

To study ray propagation in a sphere, we use the spherical coordinates (r, θ, ϕ) (Appendix 2). If θ has its origin in the radius that passes through the focus, and ϕ on a plane that contains the ray, components of the displacements of P, SV, and SH waves (Section 5.7), in the directions of the unit vectors e_r , e_θ , and e_ϕ (in the direction of increasing values of the coordinate), are (Fig. 10.3)

$$(P_r, P_\theta, 0); (SV_r, SV_\theta, 0); (0, 0, SH).$$

In terms of the angle of incidence i_0 at the observation point (Fig. 10.3),

$$\begin{aligned} P_r &= P \cos i_0; & P_\theta &= P \sin i_0 \\ SV_r &= SV \sin i_0; & SV_\theta &= SV \cos i_0. \end{aligned}$$

If the sphere represents the Earth, we have the system of geographical geocentric spherical coordinates (r, θ, ϕ) , with origins at the center of the Earth (r), North Pole of the rotation axis (θ), and Greenwich Meridian on the equatorial plane (ϕ). The components of P, SV, and SH waves in the direction of the unit vectors e_r , e_θ , and e_ϕ positive in the directions of the zenith, south, and east (Fig. 10.4) are

$$(P_r, P_\theta, P_\phi); (SV_r, SV_\theta, SV_\phi); (0, SH_\theta, SH_\phi).$$

As functions of the angle of incidence i_0 and azimuth a_z at the observation point P (Section 5.7) (Fig. 10.4),

$$\begin{aligned} P_r &= P \cos i_0; & P_\theta &= P \sin i_0 \cos a_z; & P_\phi &= P \sin i_0 \sin a_z; \\ SV_r &= SV \sin i_0; & SV_\theta &= SV \cos i_0 \cos a_z; & SV_\phi &= SV \cos i_0 \sin a_z; \\ SH_\theta &= SH \sin a_z; & SH_\phi &= SH \cos a_z. \end{aligned}$$

P and SV displacements have three components whereas SH has only two horizontal components. In relation to the geographic reference coordinate system (x, y, z) in the directions of (north, west, zenith) at the observation point defined in Section 5.7, $x = -e_\theta$, $y = -e_\phi$, and $z = e_r$.

10.2 A sphere of constant velocity

In a sphere of radius R with constant velocity V , rays propagate in straight lines and the travel time for an angular distance Δ is given by (Fig. 10.5a)

$$t = \frac{2R}{V} \sin \left(\frac{\Delta}{2} \right). \quad (10.7)$$

The travel time curve is not a straight line, as in the plane case, but a sine function (Fig. 10.5b). The ray parameter p varies with the distance Δ in the form (Fig. 10.5c)

$$p = \frac{dt}{d\Delta} = \frac{R}{V} \cos \left(\frac{\Delta}{2} \right). \quad (10.8)$$

Thus, p equals the radial distance to the center of the ray divided by the velocity (Fig. 10.5a). From this expression we obtain

$$\sin \left(\frac{\Delta}{2} \right) = \frac{1}{\eta} (\eta^2 - p^2)^{1/2}, \quad (10.9)$$

where $\eta = R/V$. Using (10.9) we derive the expression for the change of Δ with p ,

$$\frac{d\Delta}{dp} = \frac{-2}{(\eta^2 - p^2)^{1/2}}. \quad (10.10)$$

In consequence, $d\Delta/dp$ is always negative (Δ increases when p decreases) and has a singularity for $\Delta = 0$ ($p = \eta$) (Fig. 10.5d).

For a reflected wave on the free surface (Fig. 10.5a) the travel time is given by

$$t_{\text{ref}} = \frac{4R}{V} \sin \left(\frac{\Delta}{4} \right). \quad (10.11)$$

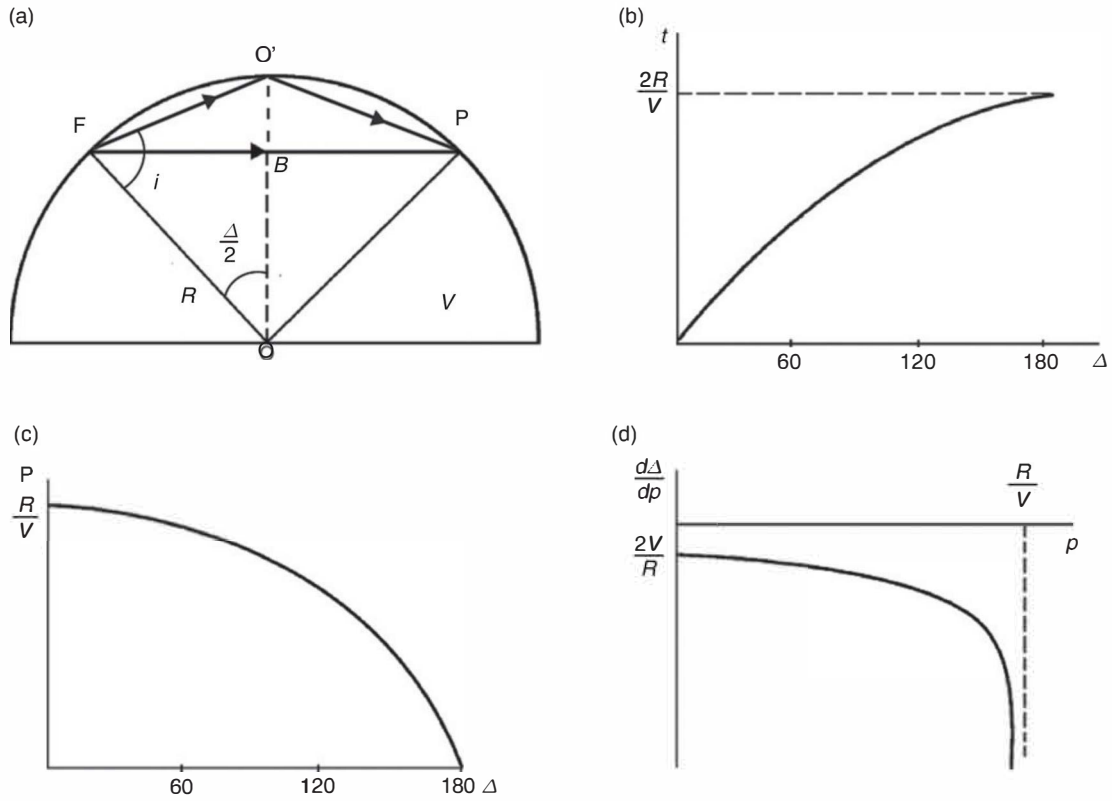


Fig. 10.5

A sphere of constant velocity: (a) a ray trajectory for a direct and reflected waves; (b) the travel time curve $t(\Delta)$; (c) a plot of the ray parameter with the angular distance $p(\Delta)$; (d) the variation of $d\Delta/dt$ with p .

10.3 A sphere with a velocity that is variable with the radius

In a sphere in which the velocity varies with the radius, as in the planar case for velocity varying with depth, the radius of curvature of the ray is given by $R_c = ds/di$ (9.7) (Fig. 10.6). Using Snell's law, we deduce

$$di = \frac{p \, dv}{\cos i}.$$

Since $ds = dr/\cos i$ (Fig. 10.7), we obtain

$$R_c = \frac{r}{p \frac{dv}{dr}}. \quad (10.12)$$

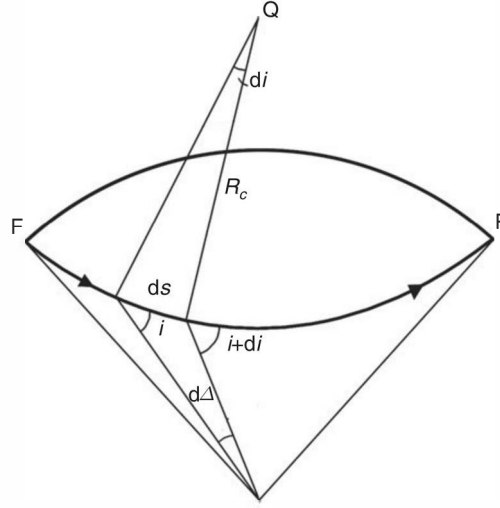


Fig. 10.6 A ray trajectory in a sphere where the velocity increases with depth and radius of curvature of the ray.

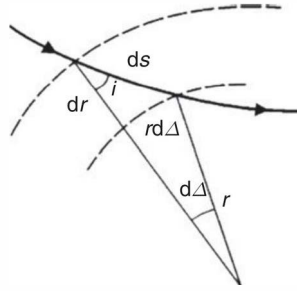


Fig. 10.7 A ray trajectory in a sphere with a velocity that depends on the radius, showing the components dr and $r d\Delta$ of an element of distance ds along the ray.

To deduce the expressions for the angular distance Δ , distance along the ray trajectory s , and travel time t , we use the geometry of the ray (Fig. 10.7):

$$\sin i = \frac{r}{R_c} \frac{d\Delta}{ds}, \quad (10.13)$$

$$ds^2 = dr^2 + r^2 d\Delta^2. \quad (10.14)$$

By substitution of (10.13) into Snell's law (10.4), we obtain

$$\frac{r^2}{v} \frac{d\Delta}{ds} = p. \quad (10.15)$$

From equation (10.14), dividing by ds^2 , using (10.15), and making the substitution $\eta = r/v$, we obtain

$$\frac{ds}{dr} = \frac{\eta}{(\eta^2 - p^2)^{1/2}}. \quad (10.16)$$

We divide (10.14) by $d\mathcal{A}^2$, and in a similar manner to what we did for the derivation of (10.16), obtain

$$\frac{d\mathcal{A}}{dr} = \frac{p}{r(\eta^2 - p^2)^{1/2}}. \quad (10.17)$$

By integrating equations (10.16) and (10.17), we obtain the expressions for the angular distance $\mathcal{A}$ and distance along the ray s . The travel time t is obtained by integration of ds/v . If the focus is at the surface we obtain

$$\mathcal{A} = 2 \int_{r_p}^{r_0} \frac{p \, dr}{r(\eta^2 - p^2)^{1/2}}, \quad (10.18)$$

$$s = 2 \int_{r_p}^{r_0} \frac{\eta \, dr}{r(\eta^2 - p^2)^{1/2}}, \quad (10.19)$$

$$t = 2 \int_{r_p}^{r_0} \frac{\eta^2 \, dr}{r(\eta^2 - p^2)^{1/2}}. \quad (10.20)$$

If the focus is at a depth h from the surface the expression is the sum of two integrals, the first from r_p to $r_0 - h$ and the second from r_p to r_0 .

In a similar way to the plane case, we can deduce the expression for the reduced time or *tau* function $\tau(p)$ (9.16), which results in

$$\tau(p) = \int_{r_p}^{r_0} \frac{1}{r} (\eta^2 - p^2)^{1/2} \, dr. \quad (10.21)$$

These expressions are very similar to those deduced for plane geometry, (9.13)–(9.16). We must bear in mind the different form of Snell's law (10.4), the use of the angular distance, and the definition of slowness.

10.3.1 The change of distance with the ray parameter

The derivation of the expression for $d\mathcal{A}/dp$ for a spherical medium is similar to that for dx/dp in plane geometry (Section 9.3). By differentiating with respect to p in (10.18), we obtain

$$\frac{d\mathcal{A}}{dp} = 2 \frac{d}{dp} \int_{r_p}^{r_0} \frac{p \, dr}{r(\eta^2 - p^2)^{1/2}}. \quad (10.22)$$

On making the substitution $f(\eta) = (1/r) dr/d\eta$ and changing the integral over r to one over η , we have

$$\frac{dA}{dp} = 2 \int_{\eta_p}^{\eta_0} \frac{f(\eta) d\eta}{(\eta^2 - p^2)^{1/2}} + 2p \frac{d}{dp} \int_{\eta_p}^{\eta_0} \frac{f(\eta) d\eta}{(\eta^2 - p^2)^{1/2}}. \quad (10.23)$$

For the second term on the right-hand side of (10.23), we integrate by parts before taking the derivative:

$$2p \frac{d}{dp} \left[f(\eta_0) \cos^{-1} \left(\frac{\eta_0}{p} \right) - \int_{\eta_p}^{\eta_0} \cos^{-1} \left(\frac{\eta}{p} \right) f'(\eta) d\eta \right]. \quad (10.24)$$

After taking the derivative, we obtain

$$\frac{-2\eta_0 f(\eta_0)}{(\eta_0^2 - p^2)^{1/2}} + 2 \int_{\eta_p}^{\eta_0} \frac{\eta f'(\eta) d\eta}{(\eta^2 - p^2)^{1/2}}. \quad (10.25)$$

By adding the first term of (10.23), we finally obtain

$$\frac{dA}{dp} = \frac{-2\eta_0 f(\eta_0)}{(\eta_0^2 - p^2)^{1/2}} + 2 \int_{\eta_p}^{\eta_0} \frac{\frac{d}{d\eta} [\eta f(\eta)] d\eta}{(\eta^2 - p^2)^{1/2}}. \quad (10.26)$$

On introducing the variable ζ related to the gradient of the velocity,

$$\zeta = \frac{r}{v} \frac{dv}{dr}, \quad (10.27)$$

it is easy to find that $\eta f(\eta) = (1 - \zeta)^{-1}$ and

$$\frac{d}{d\eta} [\eta f(\eta)] = \frac{1}{(1 - \zeta)^2} \frac{d\zeta}{d\eta}. \quad (10.28)$$

By substitution into (10.26) and changing the integral over η to one over r , we obtain

$$\frac{dA}{dp} = \frac{-2}{(1 - \zeta_0)(\eta_0^2 - p^2)^{1/2}} + 2 \int_{r_p}^{r_0} \frac{\frac{d\zeta}{dr} dr}{(1 - \zeta)^2 (\eta^2 - p^2)^{1/2}}. \quad (10.29)$$

This expression is similar to that of the plane case (9.39). On making the substitution $\xi = 2/(1 - \zeta)$, equation (10.29) can be written in the form

$$\frac{dA}{dp} = \frac{-1}{\xi_0(\eta_0^2 - p^2)^{1/2}} + \int_{r_p}^{r_0} \frac{\frac{d\xi}{dr} dr}{(\eta^2 - p^2)^{1/2}}. \quad (10.30)$$

Equations (10.29) and (10.30) are very useful when one wants to study the propagation of rays in a spherical medium for different distributions of velocity with the radius.

10.4 A velocity distribution with ζ constant

Just as in the plane case (Section 9.4), a velocity distribution with the radius in a sphere corresponding to ζ constant allows an easy derivation of expressions for the relation of the ray parameter to the distance $p(\Delta)$ and travel time $t(\Delta)$. If the velocity increases with depth (decreasing r), according to the definition of ζ (10.27), then, putting $\zeta = -\alpha$, we obtain

$$v = v_0 \left(\frac{r_0}{r} \right)^\alpha, \quad (10.31)$$

where r_0 and v_0 are the values at the surface. This distribution gives infinite velocity for the center of the sphere, so it cannot be used for very small values of r . The gradient of the velocity is negative, that is, it also increases with depth:

$$\frac{dv}{dr} = -\frac{\alpha v_0}{r_0} \left(\frac{r_0}{r} \right)^{\alpha+1}. \quad (10.32)$$

Since $d\zeta/dr$ is zero, the second term of (10.29) is null and the expression for $d\Delta/dp$ is

$$\frac{d\Delta}{dp} = \frac{-2}{(1+\alpha)(\eta^2 - p^2)^{1/2}}. \quad (10.33)$$

By integration we obtain for $\Delta(p)$ the expression

$$\Delta = \frac{2}{1+\alpha} \cos^{-1} \left(\frac{p}{\eta_0} \right), \quad (10.34)$$

and for $p(\Delta)$,

$$p = \eta_0 \cos \left(\frac{(1+\alpha)\Delta}{2} \right). \quad (10.35)$$

Since $p = dt/d\Delta$, by integration of (10.35) we obtain the expression for the travel time $t(\Delta)$:

$$t = \frac{2\eta_0}{1+\alpha} \sin \left(\frac{(1+\alpha)\Delta}{2} \right). \quad (10.36)$$

Equation (10.36) is not valid for rays passing through the center of the sphere, $(1+\alpha)\Delta/2 = \pi/2$, since there the velocity becomes infinite. This velocity distribution is known as the *Mohorovičić law* and may be used to approximate the velocity distributions for certain regions inside the Earth away from its center, for example, in the lower mantle (Section 11.5).

10.5 Rays with circular trajectory

Another special distribution of increasing velocity with depth is that corresponding to rays with circular trajectories. The radius of curvature (10.12) must be constant, which condition is satisfied if the velocity depends on the radius in the form

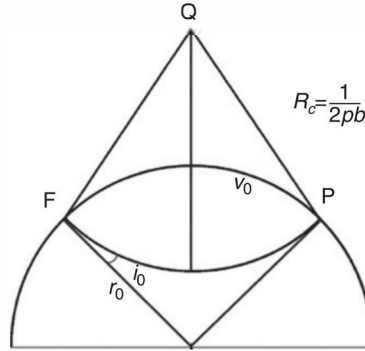


Fig. 10.8

The circular trajectory of a ray in a sphere with variable velocity $v(r)$.

$$v = a - br^2, \quad (10.37)$$

where a and b are constants. According to (10.12), the radius of curvature is now $R = 1/(2pb)$ (Fig. 10.8). In terms of the radius of the sphere and surface velocity, the velocity distribution is

$$v = v_0 + b(r_0^2 - r^2). \quad (10.38)$$

The velocity has a finite value at the center. The velocity gradient decreases with depth according to

$$\frac{dv}{dr} = -2br. \quad (10.39)$$

The radius of curvature for a ray with take-off angle at the surface i_0 , substituting $p = r_0 \sin i_0 / v_0$ (10.3), is given by

$$R_c = \frac{v_0}{2r_0 b \sin i_0}. \quad (10.40)$$

As we have seen, the velocity's distribution with the radius for rays of circular trajectories is different from that for the plane case (Section 9.5) and does not correspond to a constant gradient. It is possible to derive an analytical expression for travel time $t(\Delta)$ for this distribution, but it has a complicated form (Bullen and Bolt, 1985).

10.6 Distribution of the velocity with the radius

We have seen two distributions of the velocity with the radius inside a sphere for which it is possible to find explicit expressions, $p(\Delta)$ and $t(\Delta)$. In general this is not the case. However, we can deduce many characteristics of ray propagation for certain general distributions using the expression for $d\Delta/dp$, in a similar way to which we treated the planar case (Section 9.6). For a sphere we must remember that the slowness is defined as $\eta = r/v$ and the ray parameter is $p = \eta \sin i$. Therefore we must take into

account the changes with depth of v and r . Just as in the plane case, we will consider a distribution for which the velocity increases very slowly with depth and its gradient decreases; what is called a normal distribution. Next to this situation we add a region where the velocity increases rapidly or decreases with depth. The latter case results in the presence of a low-velocity layer. The description of the characteristics of these velocity distributions is very similar to that of [Section 9.6](#) for the plane case, so we refer the reader to that section for details. These situations with increments and decreases of velocity with depth or fast and low velocity layers are found in the structure of the Earth's interior, as we will see in [Chapter 11](#).

10.6.1 A normal distribution

In the normal distribution the velocity increases slowly with depth (decreasing r) and its gradient dv/dr , which is negative and small, changes very slowly ([Fig. 10.9a](#)). Therefore, ζ is negative (its absolute value is small and generally less than 0.25) and $d\zeta/dr$ is very small.

Ray trajectories are curved concave upward ([Fig. 10.9b](#)). The curve for $d\Delta/dp$ ([Fig. 10.9c](#)) has negative values and a singularity for $p = \eta_0$. The curve of $p(\Delta)$ exhibits a gradual decrease of p with distance ([Fig. 10.9d](#)), from a value η_0 for $\Delta = 0$. The travel time curve $t(\Delta)$ exhibits a gradual increase of time with distance with a slow decrease of its gradient ([Fig. 10.9e](#)).

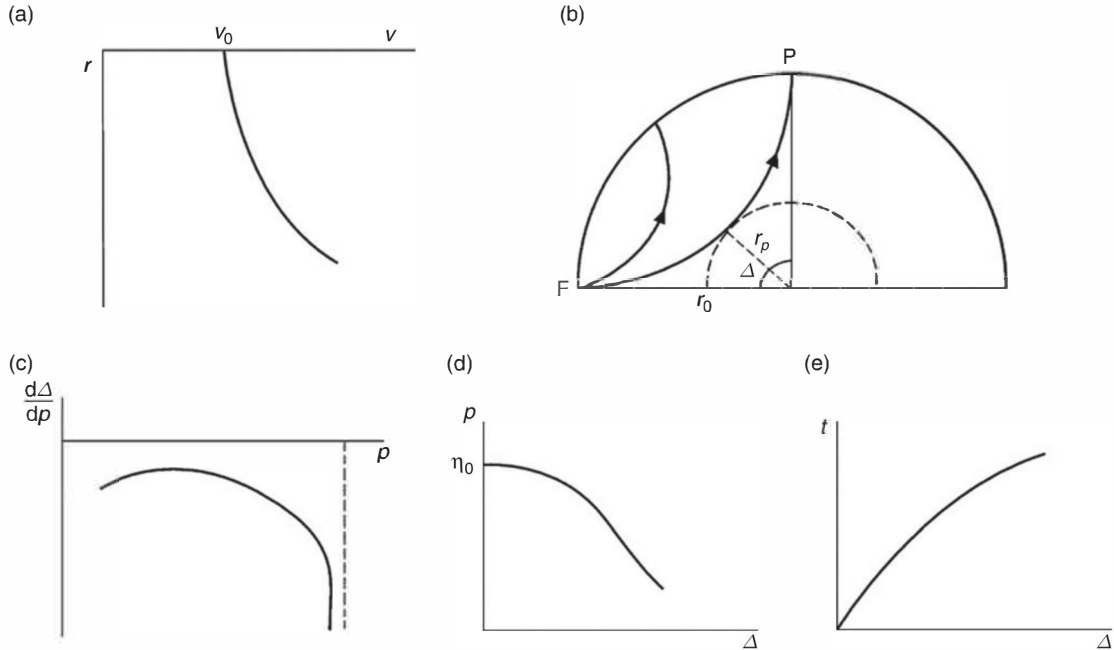


Fig. 10.9

A sphere with a normal distribution of velocity with depth: (a) the distribution of velocity with depth; (b) ray trajectories; (c) the variation of the angular distance with the ray parameter $d\Delta/dp$; (d) the variation of the ray parameter with distance $p(\Delta)$; (e) the travel time curve $t(\Delta)$.

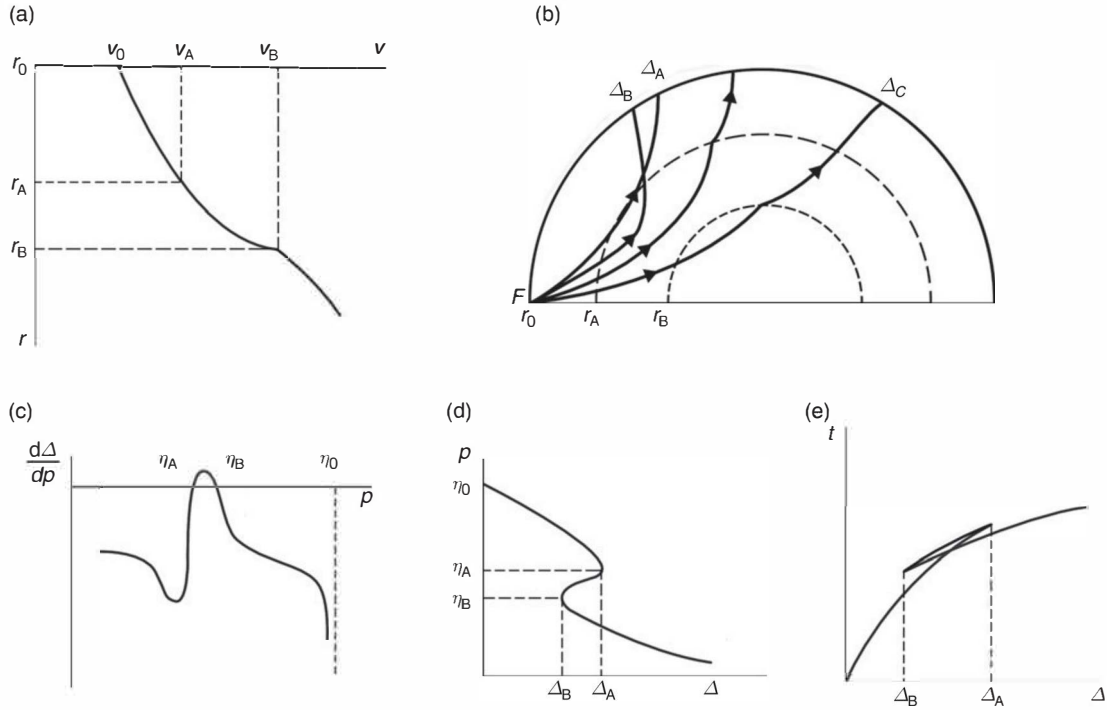


Fig. 10.10

A sphere with a zone of a rapid increase of velocity with depth: (a) the distribution of velocity with depth; (b) ray trajectories; (c) the variation of the angular distance with the ray parameter $d\Delta/dp$; (d) the variation of the ray parameter with the distance $p(\Delta)$; (e) the travel time curve $t(\Delta)$.

A velocity distribution with a small inverse power of the radius (10.31), corresponding to constant ζ , is a particular case of the normal distribution.

10.6.2 A rapid increase in velocity

Just as in the plane case (Section 9.6.2), we consider a region between the radii r_A and r_B , where the velocity increases rapidly with depth, while above and below this region the velocity distribution is normal (Fig. 10.10a). When rays penetrate into this region their radii of curvature decrease (10.12), since dv/dr is large (Fig. 10.10b). If the change in gradient is sufficiently large that the second term in (10.29) is larger than the first, then the curve of $d\Delta/dp$ has a positive part (Fig. 10.10c). In consequence, the curve of $p(\Delta)$ has a part for which p decreases with decreasing Δ (Fig. 10.10d). Then, within a certain range of values of distance, say between Δ_A and Δ_B , there are three values of p for each value of Δ , that is, three rays arrive at the same distance. Finally, the travel time curve $t(\Delta)$ (Fig. 10.10e) has a triplication with a concave upward branch that corresponds to the rays for which p decreases with decreasing Δ , with two cusps due to the caustic. These characteristics are similar to those explained for the plane case (Section 9.6.2).

10.6.3 A decrease in velocity. A low-velocity layer

As in Section 9.6.3 for the plane case, we consider a region bounded by the radii r_A and r_B , where the velocity decreases with depth, while outside of this region it increases just as it does in the normal distribution (Fig. 10.11a). In this region, rays change curvature from concave upward to concave downward due to the change in sign of dv/dr (Fig. 10.11b). According to Snell's law, rays cannot turn upward until they reach the depth where velocity has again the same value v_A as that at r_A , that is, under r_C (Fig. 10.11a). The curve of $d\Delta/dp$ has a discontinuity and the contribution of the large change in gradient leads to the fact that, for values of p between η_B and η_A , $d\Delta/dp$ becomes positive (Fig. 10.11c). The curve of $p(\Delta)$ is also discontinuous (Fig. 10.11d). Up to a distance Δ_A , rays do not penetrate into the low-velocity zone and the curve corresponds to a normal distribution of velocity. At Δ_A there is a discontinuity and, for the same value of $p = \eta_A$ a ray arrives also at distance Δ_C . Rays that arrive at distances between Δ_B and Δ_C correspond to a retrograde branch of the curve with positive values of $d\Delta/dp$, that is, p decreases with decreasing Δ (Fig. 10.11d). From Δ_B there starts a second normal branch for which p decreases with increasing Δ (Fig. 10.11d). There is, then, a range of distances $\delta\Delta$, between Δ_A and Δ_C , for which no rays arrive (there exist no values of p), namely a *shadow zone*, and another range of distances from Δ_B to Δ_C , for which two rays arrive at the same distance; arrivals at Δ_B have also a time delay δt (Fig. 10.11e).

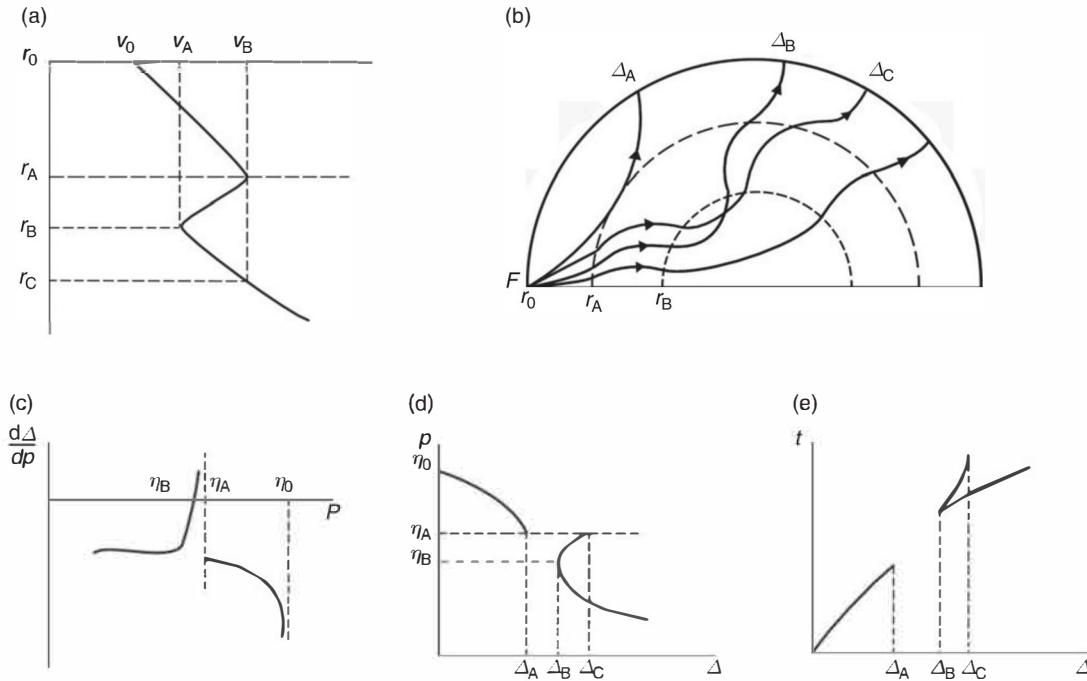


Fig. 10.11 A sphere with a zone of a decrease of velocity with depth: (a) the distribution of velocity with depth; (b) ray trajectories; (c) the variation of the angular distance with the ray parameter $d\Delta/dp$; (d) the variation of the ray parameter with the distance $p(\Delta)$; (e) the travel time curve $t(\Delta)$.

The travel time curve corresponds to these characteristics (Fig. 10.11e). From zero to Δ_A , the curve corresponds to a normal distribution. Between Δ_A and Δ_C , is the shadow zone $\delta\Delta$ where no rays arrive. Behind the shadow zone, the curve separates into two branches that come together at Δ_B with a cusp, the first with the properties of the normal distribution that continues to the limit of Δ , and the second from Δ_B to Δ_C , for which $dt/d\Delta$ decreases with decreasing Δ (a retrograde concave upward branch). Then from Δ_B to Δ_C , two rays arrive at different times at each distance. The time for Δ_B is delayed by δt with respect to that corresponding to a prolongation of the first part of the curve. This delay is due to propagation in the low-velocity layer. The main characteristics of the travel time curve are, then, the existence of a shadow zone and, behind it, a time delay and a duplication of arrivals. This situation is found in the Earth at the boundary between the mantle and the core (Section 11.6).

As we saw in Section 9.8 for the planar case, in a spherical medium we can have layers with discontinuities in velocity at their boundaries, so that reflected rays occur. Reflected rays also take place at the free surface. In the Earth this occurs in the mantle (Sections 11.4 and 11.5) and between the mantle and the core and between the outer and inner core (Section 11.6).

10.7 Determination of velocity distribution

We can derive a relation to determine the velocity distribution with the radius $v(r)$ from the travel times $t(\Delta)$ for a spherical medium in a similar form to that in the plane case (Section 9.9). We start with equation (10.18), which, on changing the integral over r to one over η , can be written as

$$\Delta = 2 \int_{\eta_p}^{\eta_0} \frac{p \frac{dr}{d\eta} d\eta}{r(\eta^2 - p^2)^{1/2}}. \quad (10.41)$$

This is an integral along a ray that arrives at a distance Δ , has a parameter p , and turns upward at $\eta = \eta_p$. Consider now all the rays that arrive at distances between $\Delta = 0$ and $\Delta = \Delta_1$, and correspond to values of p between p_0 and p_1 . For these rays we can define the integral

$$\int_{p_1}^{p_0} \frac{dp}{(p^2 - \eta_1^2)^{1/2}}.$$

By applying this operation to both sides of (10.41) we obtain

$$\int_{p_1}^{p_0} \frac{\Delta dp}{(p^2 - \eta_1^2)^{1/2}} = \int_{p_1}^{p_0} \int_{\eta_p}^{\eta_0} \frac{2p \frac{dr}{d\eta} d\eta}{r(\eta^2 - p^2)^{1/2} (p^2 - \eta_1^2)^{1/2}} dp. \quad (10.42)$$

The first integral on the right-hand side is over a generic ray of parameter $p = \eta_p$ and the second is over all rays of parameters from p_1 to p_0 , where p and η_p are variables. In the integral on the left-hand side, Δ and p are variables that take all values from 0 to Δ_1 and from p_1 to p_0 , respectively.

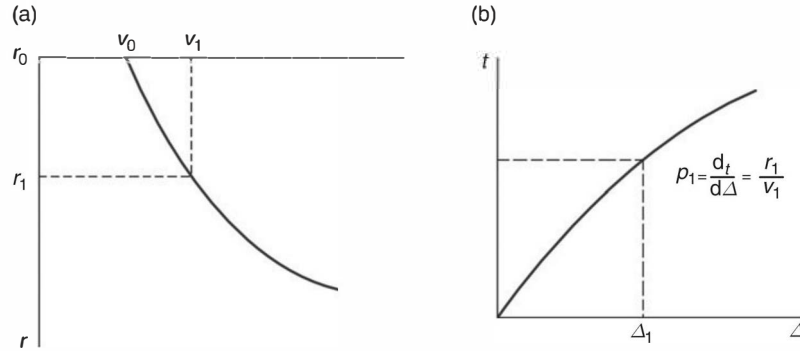


Fig. 10.12

Application of the Herglotz–Wiechert integral to obtain the distribution of velocity with depth in a spherical medium from the travel time curve. (a) The distribution of velocity with depth. (b) The travel time curve.

In the double integral on the right-hand side of (10.42), similarly to what we did in the plane case (9.63), we can change the order of integration. We integrate first over p from $p = \eta$ to $p = \eta_1$ and the result has the value of π . The integral over η from η_1 to η_0 is changed into an integral over r from r_1 to r_0 , resulting in

$$\int_{r_1}^{r_0} \int_{\eta_1}^{\eta} \frac{2p dp}{[(\eta^2 - p^2)(p^2 - \eta_1^2)]^{1/2}} \frac{dr}{r} = \pi \ln \left(\frac{r_0}{r_1} \right). \quad (10.43)$$

The term on the left-hand side of (10.42) can be integrated by parts, as we did in (9.64). Taking into account that, for $\eta = \eta_0$, $\Delta = 0$, and, for $\eta = \eta_1$, $\cosh^{-1}(1) = 0$, and changing the integral over p to one over Δ , and substituting (10.43), we finally obtain

$$\int_0^{\Delta_1} \cos^{-1} \left(\frac{p}{\eta_1} \right) d\Delta = \pi \ln \left(\frac{r_0}{r_1} \right). \quad (10.44)$$

This is the *integral equation of Herglotz–Wiechert* for a spherical medium that allows determination of the distribution of velocity with radius from the travel time. Just as in the plane case, the equation is valid only for velocity distributions in which the velocity increases monotonically with depth. It cannot be applied directly when there are rapid changes in gradient or low-velocity layers (although, as was mentioned in the plane case, there are modifications that can handle those cases).

As in the planar medium (Section 9.9), equation (10.44) is applied in the following manner. For a fixed distance Δ_1 , η_1 is known from the value of the slope of the curve $t(\Delta)$ for such a distance. To find v_1 we need to know r_1 . The value of r_1 corresponds to the turning point of the ray that arrives at a distance Δ_1 , and is obtained from equation (10.44). The integral can be calculated numerically, since, from each value of Δ , from $\Delta = 0$ to $\Delta = \Delta_1$, p is the slope of the travel time curve $t(\Delta)$ (Fig. 10.12b). For each pair of values v_1 and r_1 , we have a point of the velocity distribution (Fig. 10.12a). By repeating the process for different distances, we obtain values of the velocity for different values of r . Since usually $t(\Delta)$ are discrete observed data, we have to interpolate in order to find the values of $dt/d\Delta$. This method can be applied to obtain

velocity distributions in the lower mantle of the Earth where the conditions for validity of equation (10.44) are fairly well satisfied (Section 11.5).

10.8 Energy propagation by ray beams. Geometric spreading

The energy propagated by a ray beam inside a sphere follows the same treatment as that in the plane case (Section 9.10), with a few modifications. From the focus, an energy E is emitted homogeneously per unit time. The energy per unit solid angle is $P = dE/d\Omega$, whence $E = 4\pi P$.

If the focus is at a depth h from the surface, then the take-off angle at the focus is i_h . Rays that leave the focus between angles i_h and $i_h + di_h$ arrive at the surface at distances between Δ and $\Delta + d\Delta$, with angles of incidence between i_0 and $i_0 + di_0$ (Fig. 10.13) (if the focus is at the surface, $i_h = i_0$). The element of solid angle $d\Omega$ around the focus is

$$d\Omega = 2\pi \sin i_h di_h. \quad (10.45)$$

The area dS over the surface of the sphere defined by the intersection of the element of the solid angle $d\Omega$, at an angular distance Δ , is given (Fig. 10.13) by

$$dS = 2\pi r_0^2 \sin \Delta d\Delta. \quad (10.46)$$

The area dA defined on the wave front surface (Fig. 10.14) is $dA = dS \cos i_0$. The energy intensity, $I = dE/dA$, is the energy that arrives at a distance Δ per unit wave front surface per unit time. As a function of the energy that leaves the focus per unit solid angle, $I = P d\Omega/dA$. By substitution of (10.45) and (10.46), we obtain

$$I = \frac{P \sin i_h}{r_0^2 \cos i_0} \frac{di_h}{\sin \Delta d\Delta}. \quad (10.47)$$

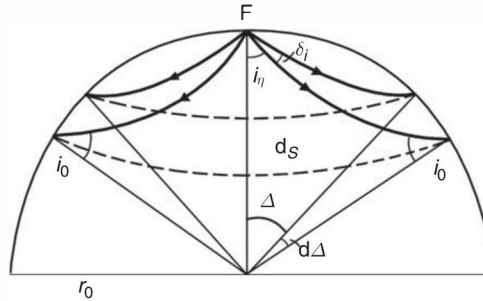


Fig. 10.13

The area on the surface of a sphere covered by the wave front corresponding to a ray beam that leaves the focus with take-off angle i_h at a distance Δ .

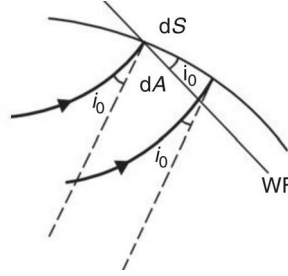


Fig. 10.14

The arrival at the surface of a sphere of a wave front corresponding to a ray beam that leaves the focus with take-off angle i_h and arrives with angle of incidence i_0 . The area of the free surface is dS . The area of the wave front surface is dA . WF: wave front.

The derivative $di_h/d\Delta$ can be expressed in terms of the slope of the travel time $t(\Delta)$ by using Snell's law:

$$\frac{d}{d\Delta} \left(\frac{r_h \sin i_h}{v_h} \right) = \frac{dp}{d\Delta} = \frac{d^2 t}{d\Delta^2}, \quad (10.48)$$

and, therefore,

$$\frac{di_h}{d\Delta} = \frac{v_h}{r_h \cos i_h} \frac{d^2 t}{d\Delta^2}. \quad (10.49)$$

By substituting (10.49) into (10.47), we find

$$I = \frac{P v_h \sin i_h}{r_h^2 \cos i_h \cos i_0 \sin \Delta} \frac{d^2 t}{d\Delta^2}. \quad (10.50)$$

If the focus is at the surface, $i_h = i_0$ and $r_h = r_0$. Hence, equations (10.47) and (10.50) are simplified.

Equations (10.47) and (10.50) represent the *geometric spreading* of energy propagated inside a sphere. The energy that arrives at a certain distance Δ decreases with the factor $1/\sin \Delta$ and depends on the value of $di/d\Delta$ or, equivalently, that of $d^2 t/d\Delta^2$. If $di/d\Delta$ is large, its inverse $d\Delta/di$ is small and Δ varies little with changes in i_h . This means that the corresponding interval of distances $d\Delta$ is small, the density of rays is large, and the energy per unit surface area is also large. The opposite happens when $di/d\Delta$ is small. Ray beams spread themselves over a larger surface so that $d\Delta$ is large and the energy density small. Because $di/d\Delta$ is proportional to $dp/d\Delta$, for distances at which $dp/d\Delta$ is large, we will have concentrations of energy. This happens when there are large changes in the velocity gradient. For those situations there are also changes in the slope of the travel time curve ($d^2 t/d\Delta^2$ is large). Hence we expect large concentrations of energy at distances corresponding to cusps of the curve $t(\Delta)$ when we have rapid increases in velocity or low velocity layers (Figures 10.10 and 10.11). The effect of geometrical spreading in the amplitude of body waves is important when we calculate synthetic wave forms (Chapter 20).

10.9 Summary

For distances greater than 1000 km in the Earth, the planar approximation is no longer valid and we must consider its spherical shape. Here we have presented ray propagation in a spherical medium beginning with one of constant velocity. Thus the derivations of [Chapter 9](#) for the planar case are now modified for a spherical medium. We began with the geometry of rays and displacements in a spherical medium with a reformulation of Snell's law (10.4) and Benndorf relation (10.6). We have applied ray theory to the spherical medium with variable velocity with radius and have derived equations for angular distance, distance along the ray, and travel times (10.18–10.20). An important relation, which gives the change of angular distance with the ray parameter in terms of the gradient of the velocity with the radius (10.29), permits us to derive the behavior of rays and the travel times in a medium with a certain distribution of velocity with radius. As examples we have presented the distribution for a ray with circular trajectory, the normal distribution in which there is a gradual increase of velocity with depth, and media in which there is a zone with a rapid increase in the gradient and a decrease of velocity with radius. Ray paths and travel times for these cases are representative of situations in the Earth interior.

For the inverse problem of the determination of the velocity distribution with radius for the spherical medium we have derived the Herglotz–Wiechert integral (10.43). Applied to the Earth this allows us to determine the structure of its interior, in the mantle and core, from the observed travel times of waves produced by earthquakes ([Chapter 11](#)). Energy propagation by ray bundles has shown a decrease of energy intensity with the inverse of the angular distance, an effect known as geometrical spreading.

10.10 Problems

- P10.1 For a spherical medium with radius R and constant velocity V , derive the equation of travel time for reflected waves (10.11) and plot the curves (t, Δ) .
- P10.2 For a spherical Earth of radius 6000 km, constant P wave velocity, $V = 10 \text{ km s}^{-1}$, for a surface focus calculate the travel times and ray parameters for $\Delta = 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$, and 180° , and draw the curves $t(\Delta)$ and $P(\Delta)$.
- P10.3 Assume that the Earth consists of two concentric regions of constant velocity: the core of radius $R/2$ and the mantle. The speed of propagation in the core is twice that of the mantle.
- Calculate the maximum angular distance of the direct ray in the mantle.
 - Plot the paths of the waves that propagate through the Earth's interior, and the travel time curves of these waves in units of R/v , where v is the speed of propagation in the mantle.

- P10.4 In a spherical Earth of radius R , the velocities of P waves have the form $v = a - br^2$. The surface velocity is v_0 and that at the center is $2v_0$. Determine the angular distance Δ for a ray that penetrates to half of the radius.
- P10.5 In a spherical medium, the velocity increases with depth according to $v = ar^{-b}$. If $v_0 = 6 \text{ km s}^{-1}$, $r_0 = 6000 \text{ km}$, and $b = 0.2$, calculate the travel times and ray parameters for $\Delta = 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$, and 180° , and draw the curves $t(\Delta)$ and $P(\Delta)$. Compare your results with the results of Problem 10.1.
- P10.6 The Earth consists of a mantle of radius R and speed of propagation $v = \frac{a}{\sqrt{r}}$, and a core of radius $R/2$ and constant speed of propagation $4v_0$, where v_0 is the speed of propagation at the Earth's surface.
- Calculate the maximum epicentral distance corresponding to a wave that travels only through the mantle.
 - Calculate the critical angle of the wave reflected at the core, and the angle at which it leaves the surface.
 - Calculate the epicentral distance Δ_c corresponding to the critical angle.
 - Plot the travel time curve, specifying Δ_c and Δ_m .
- P10.7 The Earth of radius R is formed by a mantle of velocity $v = ar^{-1/2}$ and a core of radius $R/2$ and velocity $v = br^{-1/3}$. The surface velocity is v_0 and that at the top of the core is $2v_0$. Calculate the distance Δ for a ray that leaves the surface focus with the take-off angle $i_0 = 14.5^\circ$.

In [Chapters 8 to 10](#), using ray theory, we have seen how travel time curves depend on the characteristics of the media through which the seismic waves propagate and in consequence that the velocity distribution with depth can be deduced from them. In this chapter we will apply these results to observations of wave propagation in the Earth and discuss the results concerning its internal structure. Thus we will see how we obtained the knowledge about the structure of the Earth that we introduced in [Section 2.6](#). For short distances (less than 1000 km) we can use the flat-Earth approximation and plane geometry. Seismic waves for that range of distances give information on depths of about 100 km, that is, of the crust and part of the upper mantle (lithosphere). For this range of distances we can apply the theory derived in [Chapters 8 and 9](#). For greater distances the spherical shape of the Earth must be considered, so the results of [Chapter 10](#) must be applied. The effects due to the deviations of the form of the Earth from a sphere, that is, mainly its flatness, can be taken into account by using corrections to the spherical model. In seismology these effects are not very important.

11.1 Observations and methods

The first seismic waves used for the study of the Earth's structure were those produced by earthquakes. Even today this is the main source of information, especially for the deep interior. Among the first tables and curves of travel times of seismic waves were those of Oldham, who in 1906 deduced the existence of the Earth's core. These tables were completed by Zöppritz and Herbert H. Turner and later, in 1914, by Gutenberg. In 1940, Jeffreys and Bullen published their tables of travel times which are very widely used, even today (Jeffreys and Bullen, 1940). In 1968, Eugene Herrin published tables of travel times for P waves only (Herrin, 1968). On a recommendation from the IASPEI, travel time tables have been derived from a spherically symmetric velocity model, known as *iasp91*, based on modern data (the ISC catalog 1964–88) (Kennett and Engdahl, 1991). For shorter distances, the pioneering work was done by Mohorovičić, who in 1909 discovered the discontinuity between the Earth's crust and mantle that has been given his name (shortened to *Moho*). Observations of traveling times of P and S waves generated by earthquakes continue to provide important information about the internal structure of the Earth (Bolt, 1982).

Observations of seismic waves generated by earthquakes have the limitation of lack of control over the exact place and time of their origin. For this reason, for the study of the shallowest layers of the Earth, seismic waves generated artificially by explosions and other methods are used. These studies, which were started in around 1920, are the basis of the

methods of seismic prospecting using what are known by the generic name of *seismic profiles* or *seismic soundings*. These were developed first by the oil industry and consist in the analysis of reflected and refracted waves. These measurements are called *deep seismic profiles* or *soundings* when they are applied to the determination of the structure of the Earth's crust and upper mantle. They are divided into two types, namely, *refraction* and *wide-angle reflection* and *vertical reflection* (Prodelh and Mooney, 2012). The first method uses waves refracted and reflected with large angles of incidence at long distances and the second uses waves reflected vertically at very short distances. In both cases, information consists in traveling times and amplitudes of waves observed along linear profiles, but quite different treatments are applied to the data.

Another, more recent, method developed to study the interior of the Earth is that of *seismic tomography*. This method, based on techniques borrowed from other fields, especially medicine, allows the mapping of inhomogeneities existing in the medium traversed by seismic waves. The method consists in the observation of a very large number of seismic rays that cross a given region in different directions. Applications of this method have resulted in the establishment of three-dimensional models of the Earth's interior. Information about the structure of the crust can be also obtained by the method known as *receiver functions*.

In recent years different projects, such as EARTHSCOPE (USA) and EPOS (EU), have been established to improve our knowledge of the structure of the Earth's interior and its evolution. In these projects large surveys are carried out using thousands of seismic, GPS, and other types of geophysical instruments ([/ds.iris.edu/ds/nodes/dmc/earthscope/](https://ds.iris.edu/ds/nodes/dmc/earthscope/); www.epos-ip.org/; last accessed May 2016).

The first complete velocity models of the Earth's interior based on seismological observations were developed in around 1930 by Bullen, Gutenberg, Jeffreys, Macelwane, and Richter, among others. These models have radial symmetry separating the Earth's interior into three regions: the crust (0–30 km), mantle (30–2900 km), and core (2900–6370 km). They have very similar general characteristics, although they differ in terms of the transition zones, especially those between the mantle and the core and between the outer and inner cores. The widely used *Jeffreys–Bullen (JB) model* was later recognized to have an average offset of 1.8 s in the travel times. More accurate velocity models with spherical symmetry have been presented. They include the *Preliminary Earth Model (PREM)* (Dziewonski and Anderson, 1981) and *CAL8* (Bolt and Uhrhammer, 1981) (Fig. 11.1, Appendix 6). More refined models using the ISC catalog for 1964–88 with a better fit to the observed travel time data have recently been developed, namely, the already mentioned *iasp91* (Kennett and Engdahl, 1991), which has been improved for the core phases by model *ak135* (Kennett *et al.*, 1995), and model *SP6* with slightly higher velocities in the upper mantle than *iasp91* (Morelli and Dziewonski, 1993). These models have radial symmetry and are commonly used as reference models for the study of lateral inhomogeneities. Actually, more detailed velocity models of the Earth's interior are three-dimensional (3-D) and include also inelastic properties and anisotropy (see Chapter 15) (Boschi *et al.*, 1996, Romanowicz and Dziewonski (eds.), 2015). Present heterogeneous 3-D seismic velocity models allow more accurate calculations of travel times and wave forms (Igel *et al.*, 2000, Takeuchi *et al.*, 2000). Numerical values of travel times for

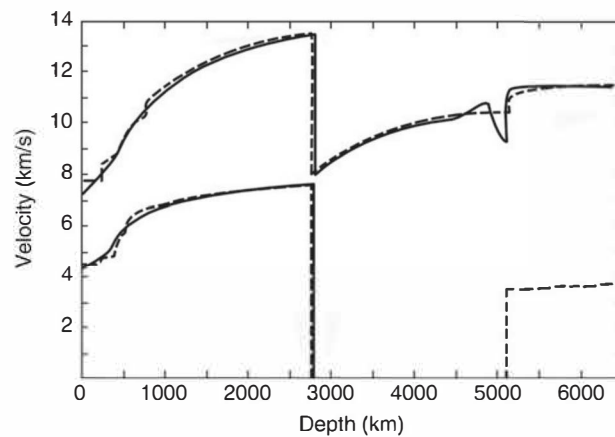


Fig. 11.1

Velocity distribution of P and S waves inside the Earth according to the Jeffreys–Bullen (continuous line) and PREM (dashed line) models (modified from Bullen and Bolt (1985)).

different Earth models may be obtained using, for example, the *TauP tools* (<https://ds.iris.edu/ds/nodes/dmc/software/downloads/>; last accessed May 2016).

11.1.1 Refraction and wide-angle reflection

The first methods to study the structure of the Earth's crust and upper mantle based on the observation of seismic waves generated by explosions are those using refraction and wide-angle reflection profiles. These methods are based on those in use in the oil industry during the 1920s. Their use for the study of the structure of the crust has been very extensive, since approximately 1940, especially after the development of portable seismograph stations with magnetic recording in 1960. Since the penetration of seismic rays depends on the distance of observations, long profiles are used to reach depths down to the base of the crust and layers of the upper mantle. This implies the use of large explosions and recordings at great distances. The size of explosions establishes the practical limit on the depth that can be investigated with these methods. Environmental regulations about explosions in the sea have led to the use of new methods, such as the *streamer method*, which uses as sources several arrays of air guns of different volumes tuned to obtain an impulse source.

The general outline of these methods consists in recording inland by portable seismographic stations along lines of length from tens of kilometers to nearly 1000 km at intervals of 1–5 km, of seismic waves generated from explosions whose size may vary from tens of kilograms up to five metric tons of explosives (Fig. 11.2). Explosions are generated in holes drilled in land or by use of underwater charges at various depths (60–200 m) in the sea. In marine profiles, receivers (hydrophones usually spaced at 12.5 m) are mounted on a neutrally buoyant cable up to 9 km long which is towed at a fixed depth and offset behind the vessel and the source.

Basically, these methods are based on the analysis of travel times and amplitudes of recorded waves according to the theory developed in Chapters 8 and 9, for layered media of

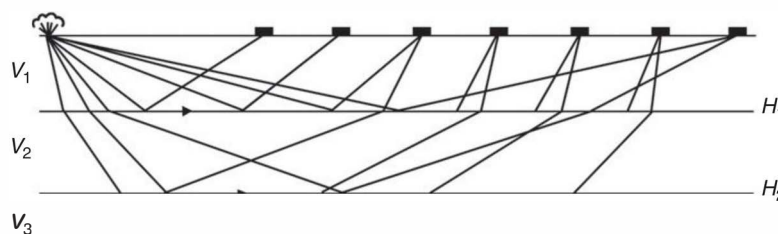


Fig. 11.2

An explosion and the line of recording stations for a refraction and wide-angle reflection seismic profile in a two-layer medium.

constant velocities and continuous distributions of velocity with depth with different velocity gradients. The forms of travel time curves, critical distances of reflected waves, and slopes of refracted waves, together with distributions of amplitudes, allow the determination of velocity models for the Earth's crust. When possible, profiles are recorded in both directions along the same line, with so-called *direct* and *inverse profiles*, in order to detect dipping layers and true velocities (Section 8.5). By obtaining profiles with a high density of recording stations, detailed models with lateral heterogeneity of layer thicknesses and velocities can be obtained.

A very common representation of observations in seismic profiles is that of *record sections*, whereby complete seismograms are represented versus distance instead of only travel times, so that the information includes also amplitudes of waves. Examples of record sections are shown later in Fig. 11.8 for a theoretical model of the crust and in Fig. 11.11 for an observed profile. A technique of interpretation is the comparison of observed and theoretical record sections calculated by the method of *ray tracing* (see Fig. 11.8, later). This method provides synthetic seismograms at given distances from models of the crust with velocity gradients and lateral changes in velocity and layer thicknesses. Seismic refraction and wide-angle reflection profiles have provided abundant observations on the structure of the crust and upper mantle for the past 60 years.

11.1.2 Vertical reflection

Since 1950, workers involved in industrial seismic exploration have abandoned methods of seismic refraction in favor of vertical reflection. These methods are based on observation of waves reflected almost vertically at short distances from the source. The source of energy can be small explosions or mechanical vibrators coupled to the surface. These methods allow one to identify with great detail, using complex processes of data reduction, the presence of reflectors at various depths, and provide an accurate image of the structures present underground (Sheriff and Geldart, 1982). In around 1965, these methods, developed by the oil industry, started to be used for the study of the crust in order to obtain clear reflections from its base and those of intermediate layers. Since 1970, many programs of extensive deep vertical reflections have been started in many countries, for example COCORP in the USA, BIRPS in the UK, ECORS in France, DEKORP in Germany, and IBERCORS in Spain, some of which continue to the present day.

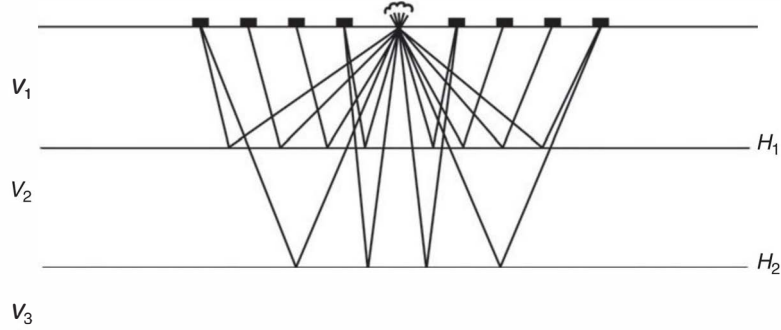


Fig. 11.3

An explosion and the recording stations for a vertical reflection seismic profile in a two-layer medium.

Vertical seismic reflection is based on the generation of reflections in layers of the crust's interior by rays of near vertical incidence using short distances from the source to observing stations (Fig. 11.3). Assuming that the crust is formed by layers of constant velocity, the equations for the times and distances of reflected waves are given by (8.68) and (8.69). For vertical reflections, the traveling time of a reflection at layer k is

$$t_k = 2 \sum_{i=1}^{k-1} \frac{H_i}{v_i}. \quad (11.1)$$

This is called the *two-way time* for normal incidence at layer k , that is, the time taken for the wave to travel from the surface to the top of layer k and back to the surface. Layers with clear reflections are called *reflectors*. The depth H , at which each reflector is located, is the sum of the thicknesses H_i of all the layers on top of the reflector. Since the observation is t_k , to obtain H we need to know the velocities v_i of all layers on top, or at least a mean value of the velocity from the surface to the reflector. Owing to the difficulty of determining velocities, depths to reflectors are often expressed in terms of two-way times.

To increase the signal-to-noise ratio of reflections from a given reflector, techniques developed by the oil industry are used. A common technique is the *stacking* or sum of amplitudes of several near records into a single trace. This method is in common with other seismological studies, such as phase identification and source directivity studies (Udías *et al.*, 2014). Each record is corrected for the differences among two-way times due to the differences in distance by using what is called the *normal move out* (NMO) correction, given by

$$\Delta t = \frac{x^2}{2v^2t}, \quad (11.2)$$

where x is the distance from each trace to the point where stacking is performed and t is the two-way time for that point. This correction permits the summation of several amplitudes as if they were located at a single point and produced by a common point of reflection (the *common depth point*; CDP). Owing to the fact that layers are not flat but rather have different dips, vertical displacements, and other types of irregularities, the traveling times of

reflected waves do not have the simple form of equation (11.1). Besides, we have the presence of diffracted waves and multiple reflections. To eliminate the influences on recordings of these undesirable effects, which may conceal the real position of reflectors, a technique called *migration* is used. There are many methods for performing migration that involve, in general, complex data processing. The results are *migrated sections* from which diffractions and multiple reflections have been removed and in which non-flat reflectors are correctly located.

Another problem to consider in the resolution of vertical reflections is the influence of the frequency of signals. As we saw in Chapter 8, ray theory, on which all these methods (refraction and reflection) are based, is valid only for high frequencies. There is, then, a limit to the resolution of lateral dimensions or vertical irregularities of reflectors that can be detected with waves of a certain frequency. Approximately, for vertical dimensions to be detectable they must be greater than a quarter of the wave length of the signal. For example, for waves of 20 Hz and velocity 6 km s^{-1} , the minimum size of detectable bodies is about 70 m. For horizontal dimensions of reflectors, we must also take into account the *Fresnel zone* (after Augustin-Jean Fresnel), that is, the zone from which energy is reflected by half a cycle of a spherical wave incident on a plane surface. Inside the crust, for waves of 20 Hz, minimum detectable horizontal dimensions are of the order of kilometers. Vertical reflection profiles allow correlation of reflectors to the continuation in depth of geologic structures at the surface. Even though this method does not give distributions of velocities, taken together with refraction profiles, it has contributed to the establishment of detailed models of the crustal structure.

11.1.3 Seismic tomography

The method of seismic tomography applied to the study of the Earth's interior was developed from the mid 1970s (Aki *et al.*, 1976). This method is based on study of the time residuals of the arrival of seismic waves with respect to those expected from an initial model. The foundations of the general theory of tomography can be found in the work of the mathematician Johann Radon in 1917, and its first applications developed by Allan M. Cormack and Godfrey N. Hounsfield between 1960 and 1970. Tomography has been applied to other fields, for example in medicine, to obtain detailed images of the human body from X-rays crossing in many directions. To study the Earth's interior, a large number of seismic rays that cross a particular region in many directions are used. From observations of the arrival times of these rays, we can find variations in velocity in the studied region with respect to a reference model. In this way, the method permits the deduction of three-dimensional models of inhomogeneities in the Earth's interior from the crust to the core, such as have been shown to exist by the work of Robert W. Clayton, Adam Dziewonski, Don Anderson, Ichiro Nakanishi, and Guust Nolet, among others (Nolet (ed.), 1987, Iyer and Hirahara, 1993, Rawlinson and Sambridge, 2003).

A commonly used method of seismic tomography is based on time residuals of arrival times, that is, the differences between observed travel times and those calculated from a reference model, usually with a distribution of velocity with depth ($\Delta t = t_{\text{obs}} - t_{\text{theo}}$). Time residuals are interpreted in terms of anomalous structures of velocity with respect to the

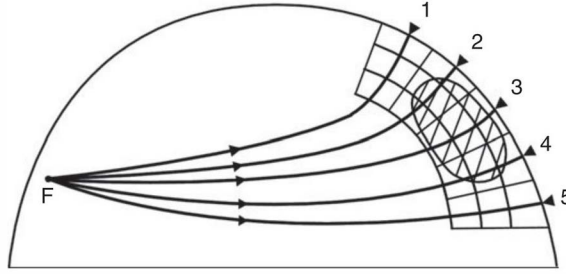


Fig. 11.4

Ray paths that cross an anomalous region (shaded) inside the Earth in a study of seismic tomography. Stations 2, 3, and 4 present traveling time residuals, whereas stations 1 and 5 do not. Cells crossed by rays are shown.

reference velocity model. Reference models of the Earth are taken from those already established with distributions of velocity with the radius (JB, PREM, etc.). If the observed traveling time at one station differs from that of the reference model for the same trajectory, the residual is assigned to a difference in velocity in some region along the ray. The aim of tomography is to explain residuals of travel times in terms of variations in the velocity distribution inside the Earth or velocity anomalies. If we have a large number of ray trajectories in many directions, the anomalous regions can be defined well in terms of the extent and the magnitude of the anomaly.

Since the travel times of waves depend on the inverse of the velocity (the slowness, $\eta = 1/v$), this variable is the one that is generally used. If the anomaly in velocity in a certain region is $\Delta v = v' - v$, the anomaly in slowness is

$$\Delta\eta = \frac{1}{v + \Delta v} - \frac{1}{v} \simeq -\frac{\Delta v}{v^2}. \quad (11.3)$$

Once the time residuals Δt have been corrected for other factors, they can be assigned to anomalies in velocity (slowness) in a given region (Fig. 11.4). If the length of a ray's trajectory through the anomalous region is l , then

$$\Delta t = l\Delta\eta = -\frac{l\Delta v}{v^2}. \quad (11.4)$$

In general, the region of the Earth that is being studied is divided into cells of constant dimensions, with a velocity that differs from that of the reference model by a quantity Δv , which must be determined. The time residual Δt at a station with respect to the reference model is assigned to the sum of time increments, positive or negative, due to velocity anomalies in each cell along the path:

$$\Delta t = \sum_{i=1}^N \frac{l_i \Delta v_i}{v_i^2}. \quad (11.5)$$

If a particular ray crosses N cells where the velocities of the reference model are v_i , the relation between Δt and Δv_i is given by equation (11.5). The length of the ray path across the cells can be written in terms of the vertical dimensions of each cell h , which usually are

the same for all cells, and the angles of incidence of the ray at those points, $l_i = h/\cos i_i$. Equation (11.5) has N unknowns (Δv_i), the velocity anomalies of cells crossed by one ray. The complete model has a total number of M cells, of which each ray crosses only a limited number N ($N < M$). The problem consists in finding Δv_i for each of the M cells. This is possible only if we have many rays crossing each of the M cells in different directions. The problem requires solution of a system of many linear equations of the type (11.5) for the M unknowns (Δv_i). The solution is generally found by a least squares method. For the solution to be well conditioned, the total number of equations must be much larger than M and each cell must be crossed in different directions by many rays. This implies that we must have a dense distribution of stations that receive rays from a large number of sources. Owing to the large number of observations, complex inversion algorithms and statistical methods are used.

This explanation has shown only the fundamentals of the method, without mentioning many other important aspects, such as the influence of observation errors, errors in hypocenters, and ray trajectories. All of these must be considered if one is to obtain results that are representative of the velocity structure of the region being studied. Seismic tomography studies provide the distribution of zones of positive or negative velocity anomalies with respect to those of the reference model of the Earth. Positive anomalies represent zones of higher velocity and are associated, considering the material homogeneous, in terms of temperature variations with regions where the material is more rigid and colder; whereas negative anomalies corresponding to lower velocities are associated with less rigid and warmer material.

11.1.4 Receiver functions

Another method for studying the crust and upper mantle structures is *receiver functions* analysis (Langston, 1977; Kind *et al.*, 2002; Alinghi, 2003). The basic idea of this method is the analysis of P and S waves incoming from teleseismic distant sources. Body waves recorded at a teleseismic station contain information on the source, propagation effects of the media (crust and mantle), and local structures beneath the station. If we remove the source and the path effects, we can obtain information on the structure under the station. When P waves arrive at a discontinuity, for example at the base of the crust under the station (Fig. 11.5), they can continue as P (Pp), or be converted to S (Sp) waves, and in the case of S continue as S (Ss) or be converted to P (Sp). As S waves travel slower than the P waves, if we measure the time difference between the two arrivals (Pp and Ps), we can obtain information about the depth of the discontinuity, assuming a velocity model for the region. Other phases corresponding to different conversions on the discontinuities can also be used. Usually observations from several near stations are grouped into clusters that sample the same structure.

To obtain the so called *receiver functions*, we must first remove the instrument from the three components (Z, NS, and EW) and rotate it to obtain the component along the ray (the longitudinal, r) (P), the Q (normal to the r component in the vertical plane) (SV), and the transverse components (SH). Usually observations are grouped by azimuth and distance and equalized to a common distance, and move-out corrections are made. Finally, by

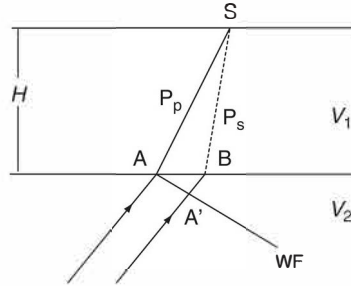


Fig. 11.5 P wave conversion into Pp and Ps arriving at station S. WF: wave front.

deconvolution, on time or frequency domain of the Q component with P signal on the r component, we remove the source and path effects. To obtain the structure under the station we generate synthetic receiver functions using different models and compare them with the observed ones. The fit between observed and synthetic receiver functions allows us to obtain the presence and location of discontinuities for a layered model beneath the station.

11.2 Distribution of velocity, elasticity coefficients, and density

The methods we have briefly described provide knowledge of velocity distributions of body waves inside the Earth. In most cases these methods use P waves only, and the Earth's structure is given as distributions of their velocity. Observations of S waves are less commonly used and the distributions of their velocity are less accurate. This is due, in part, to the fact that S waves are not first arrivals and they are not efficiently generated by explosions. If the velocities α and β are known, we can deduce the distribution of Poisson's ratio σ . The quotient β/α can be expressed in terms of σ by using (4.28), (4.72), and (4.73):

$$\frac{\beta}{\alpha} = \left(\frac{2\sigma - 1}{2(\sigma + 1)} \right)^{1/2}. \quad (11.6)$$

Also, σ can be written in terms of α and β as

$$\sigma = \frac{\alpha^2 - 2\beta^2}{2(\alpha^2 - \beta^2)}. \quad (11.7)$$

In this form, if we know independently the distributions of α and β with depth, we can deduce that of σ .

The elastic coefficients μ (the rigidity modulus), K (the bulk modulus), and λ (the Lamé coefficient) can be expressed in terms of the velocities of elastic waves α and β , and the density ρ :

$$\mu = \rho \beta^2, \quad (11.8)$$

$$K = \rho \left(\alpha^2 - \frac{4}{3} \beta^2 \right), \quad (11.9)$$

$$\lambda = \rho (\alpha^2 - 2\beta^2). \quad (11.10)$$

The variation of density with depth inside the Earth may be determined from the velocity of seismic waves α and β under certain conditions. In a simplified form, let us assume that a material is homogeneous, in hydrostatic equilibrium, and existing under adiabatic conditions. In this case, the variation of pressure P with the radius inside the Earth is

$$\frac{dP}{dr} = -g\rho, \quad (11.11)$$

where g is the gravitational attraction of the mass inside a given value of r (for a spherical Earth, $g = -Gm/r^2$). The variation of density ρ with the radius for hydrostatic equilibrium is

$$\frac{d\rho}{dr} = \frac{d\rho}{dP} \frac{dP}{dr}. \quad (11.12)$$

Since, by the definition of K , for a homogeneous material at constant volume and adiabatic change (4.33),

$$\frac{1}{K} = \frac{1}{\rho} \frac{d\rho}{dP}. \quad (11.13)$$

By substituting (11.11) and (11.13) into (11.12) and expressing K in terms of α and β (11.9), we obtain

$$\frac{d\rho}{dr} = \frac{Gm\rho}{r^2 \left(\alpha^2 - \frac{4}{3} \beta^2 \right)}. \quad (11.14)$$

This expression is known as the *Adams–Williamson equation* (after Leason H. Adams and Erskine D. Williamson). The gradient of the density with the radius for a value of r can be determined from values of α and β and the mass of the Earth inside such a value of r . This equation can be applied to obtain the distribution of density inside the Earth, if there are no discontinuities in the values of ρ and consequently also none in those of α and β . Once the distribution of ρ is known, we can find those of the elastic coefficients K and μ , from values of α and β , according to (11.9) and (11.8). For a more general treatment see Bullen and Bolt (1985).

The distributions of μ , K , and ρ inside the Earth are shown in Fig. 11.6. We can see that the density increases gradually with depth from 3.4 g cm^{-3} in the mantle and there is a sharp increase from 5.57 g cm^{-3} in the base of the mantle to $10\text{--}13 \text{ g cm}^{-3}$ in the core. Owing to this high density, the largest contribution to gravity is from the mass of the core.

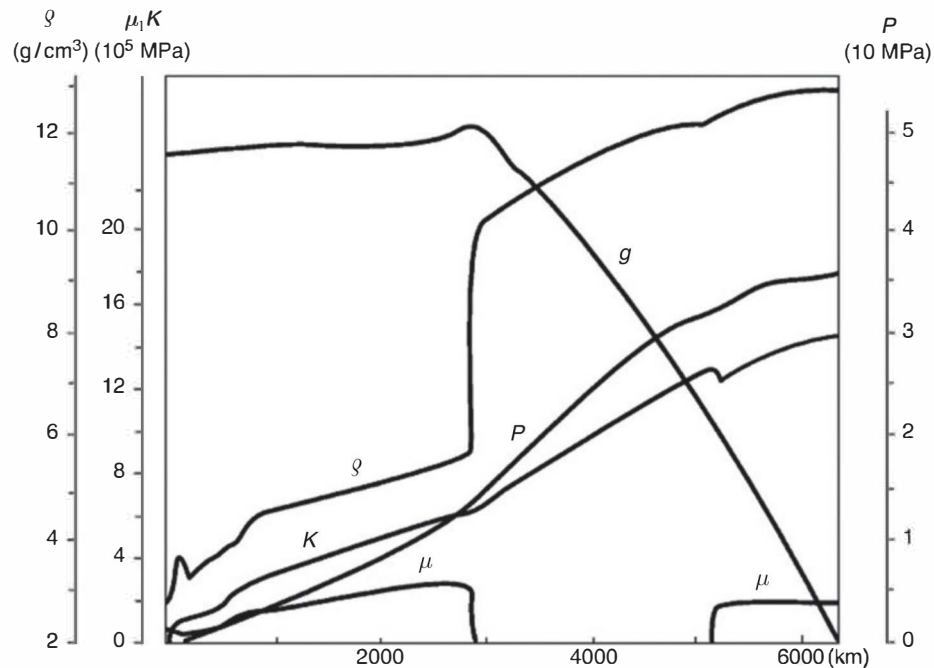


Fig. 11.6 Distributions in the Earth's interior of gravity (g), density (ρ), pressure (P), bulk modulus (K), and rigidity modulus (μ).

The pressure increases gradually with depth in the mantle and core. The bulk modulus K increases slowly in the mantle and more rapidly in the core. The rigidity modulus μ increases very slowly in the mantle, is null inside the outer core, which is in a fluid state, and has a finite value in the inner core.

11.3 The crust

As a first approximation the crust of the Earth can be considered to be formed by two layers (upper and lower crust) on top of the mantle (Section 2.6), overlaid by a thin layer of sediments. The nomenclature used to designate rays traveling through the crust is the following: rays in the upper crust are denoted P_g , waves reflected in the Moho are denoted $P_M P$, and waves refracted in the upper mantle are denoted P_n . Rays refracted and reflected in the internal discontinuity of the crust (the discontinuity between the upper and lower crust) are P_1 and $P_1 P$, but they are very difficult to observe. Earthquake waves refracted in the lower crust are also called P^* . For S waves the nomenclature is similar, that is, S_g , S_n , S^* , S_M , S , etc. For reflected waves we must also consider converted waves, that is, incident P or S waves that are reflected as S or P waves. For rays reflected in the Moho, they are denoted as $P_M S$ and $S_M P$.

Rays traveling in a two-layer crust with a sedimentary layer are shown in Fig. 11.8, with the corresponding theoretical reduced record section (reduced velocity 6 km s^{-1}) (Sections 8.4 and 9.8). For distances less than 140 km, first arrivals correspond to P_g and, for greater distances, to P_n , which have smaller amplitudes (Fig. 11.9). Depending on the thickness of the crust, this distance may vary in the range 90–160 km. A prominent arrival belongs to $P_M P$, with large amplitudes, especially near the critical distance (80–140 km). The critical distance of $P_M P$ can be used to determine the total thickness of the crust (Section 8.4). In seismograms of local earthquakes, we observe the same phases but P_g is observed with large amplitudes beyond the critical distance and S phases have large amplitudes in the horizontal components (Fig. 11.10). Rays refracted or reflected at internal discontinuities of the crust, P_I or P^* and S_I or S^* , have, in general, small amplitudes. They are observed better in seismograms of earthquakes than they are in seismic profiles.

Low-velocity layers have often been observed in the crust. Their presence in the upper and lower crust produces delays in rays reflected and refracted at discontinuities below them. In general, models of the crust with layers of constant velocity reproduce many observed characteristics of travel times and record sections. Better results are obtained upon introducing velocity gradients in the layers. A simple model of this type is that shown in Fig. 11.8, with two layers with small gradients. Rays have curved trajectories and there is a maximum distance for those traveling inside the upper crust (P_g) and those reflected in the mantle ($P_M P$). The real situation is, in general, more complex than those predicted by simple models, as can be seen from Fig. 11.11. The main phases, however, correspond, in general, to those of the synthetic example of the two-layer model (Fig. 11.8).

In oceanic regions, the structure of the crust is different from that explained for a stable continental crust. First of all, its thickness is smaller, 6–10 km, and it is formed by a thin

	oceanic	cont.	orogen.
0	1	4.5	4
	4.5		
10	6.5	6.0	6.0
20	7.8–8	6.5	
30		6.5–7	6.4
40		8–8.2	6–6.7
50			8–8.2

Fig. 11.7 Typical values of P wave velocity (km s^{-1}) in oceanic material, continental shield, and orogenic crust.

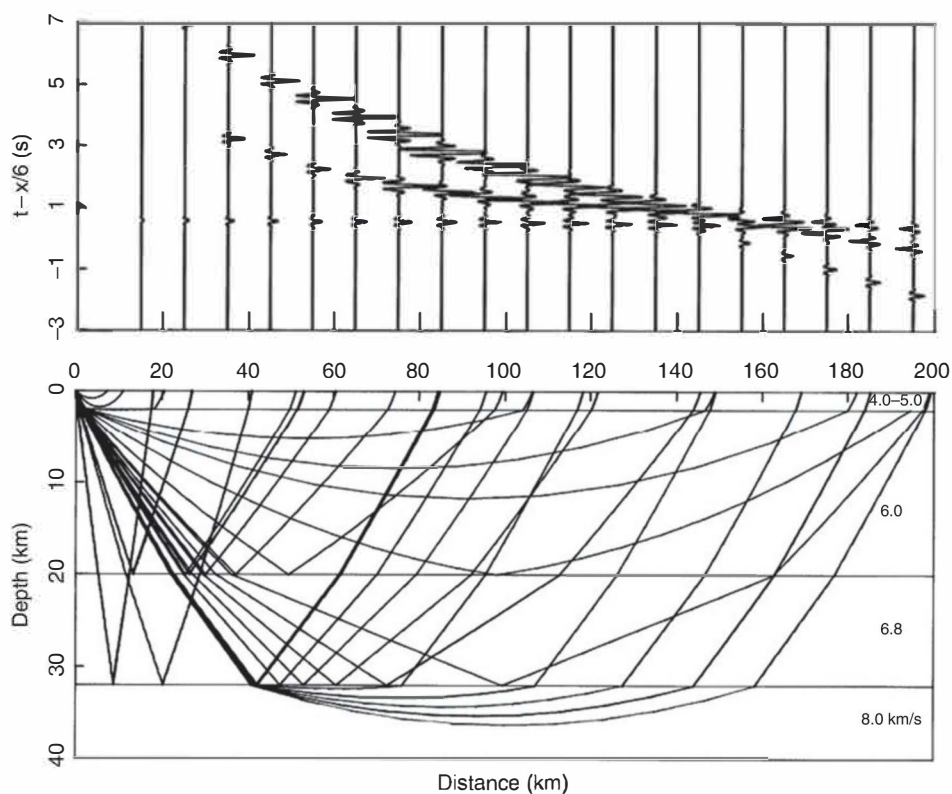


Fig. 11.8 Ray trajectories and a synthetic seismic record section for a two-layer crust with small velocity gradients (courtesy of J. Tellez).

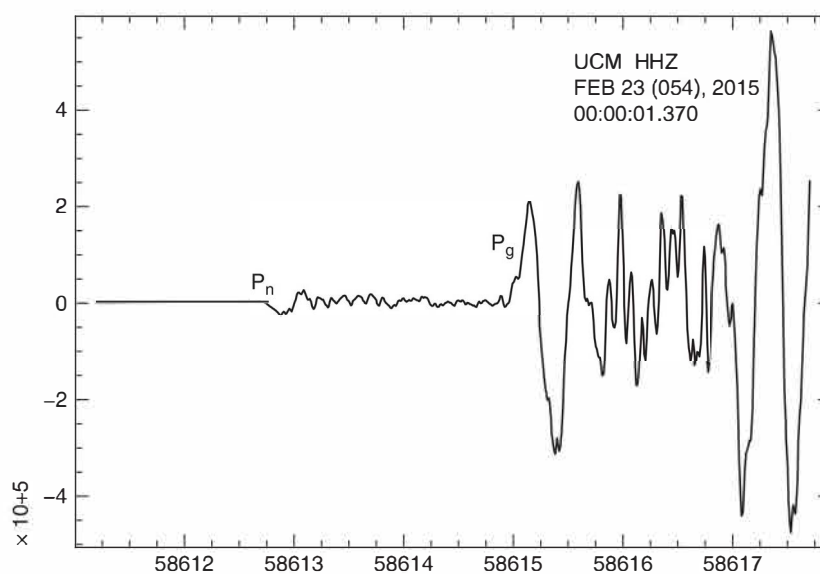


Fig. 11.9 Arrivals of P_n and P_g at station UCM (Madrid, Spain) for an earthquake at epicentral distance 150 km.

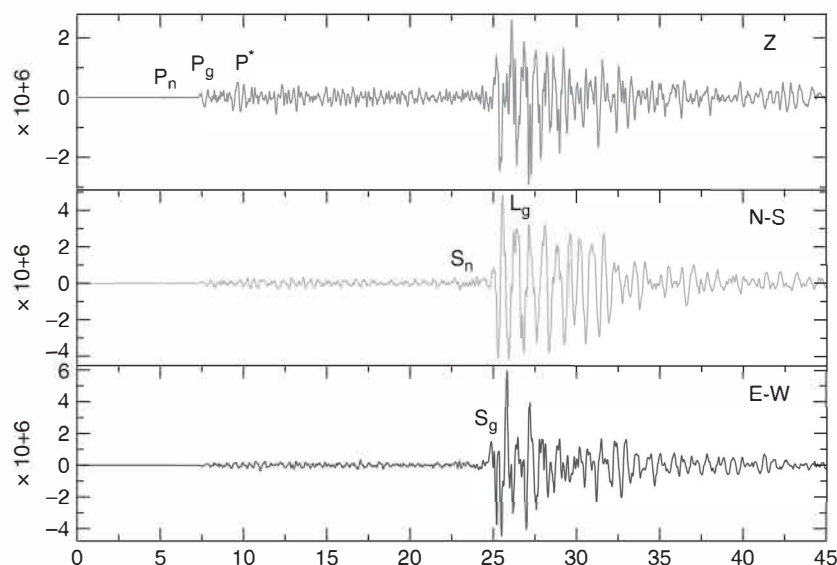


Fig. 11.10 A seismogram of a local earthquake in Ossa de Montiel, Spain (23 February 2015), recorded at the UCM BB station (Madrid, Spain) at 150 km distance, showing crustal phases. (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

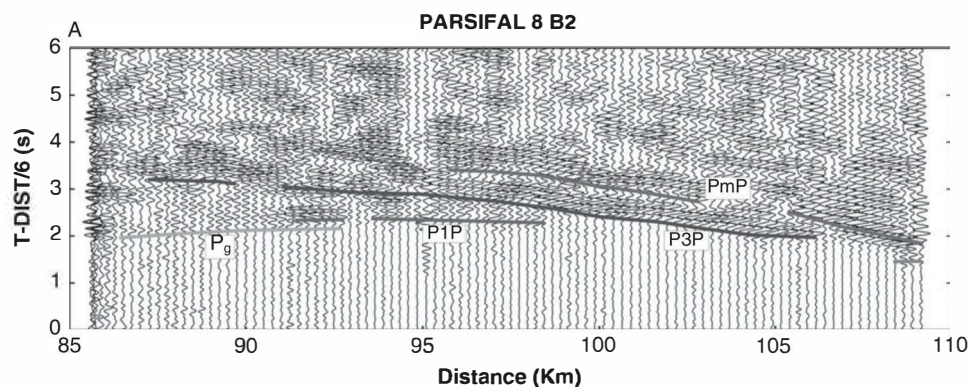


Fig. 11.11 An observed seismic record section of a refraction seismic profile offshore in the Gulf of Cadiz, Spain. P_g , P_1P , P_3P , and P_mP phases are clearly recorded; P_1P and P_3P are reflections at intermediate crustal layers (courtesy of T. Medialdea). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

layer of sediments underlain by a single layer with P velocities in the range $6.6\text{--}7\text{ km s}^{-1}$, corresponding to what has been called the basaltic layer or lower crust under continents. Underneath, in the upper mantle, P wave velocities are in general lower than they are in continental regions, $7.6\text{--}7.9\text{ km s}^{-1}$ (Fig. 11.7). This is manifested by smaller travel time values for critical distances of P_mP waves and by there being lower slopes of rays refracted inside the crust than those for continents. This holds for a typical oceanic crust

corresponding to oceanic abyssal planes (of about 4000 m depth). For shallow seas and near the coast, the crust is thicker and velocities are a little lower.

The structure of the crust below continents is not homogeneous. In regions of mountain ranges, also called *orogenic zones*, the thickness is usually larger than that observed for shields or stable regions. In zones of large mountains, such as in the Andes and Himalayas, the crustal thickness may be as great as 60 km, that is, nearly double that of stable regions. The first seismological observations that revealed this effect were due to Gutenberg for the Alps and Perry Byerly for the Sierra Nevada, California. This thickening of the crust or roots under mountains agrees with the *isostatic models* proposed by George B. Airy in 1856. The structure of the crust in mountain regions may differ also from that of stable regions with more homogeneous characteristics, and in some cases there is no discontinuity between the upper and lower crust. The transition from continental crust to oceanic crust in coastal regions is gradual, with decreasing thicknesses and disappearance of the granitic layer. The surface of the Moho has, in consequence, a complex topography that resembles an inverse image of the free-surface topography (Fig. 11.12).

The crustal structure, which is relatively constant in continental shields or stable regions, varies rapidly from one place to another in geologically complex regions (Taylor and McLennan, 1985; Meissner, 1986; Worthington *et al.*, 2016). Laterally homogeneous models of the crust are, then, not very representative. The base of the crust and the intermediate layers inside have dips and gradients that vary from place to place. Detailed refraction studies have revealed these lateral inhomogeneities. Vertical reflection surveys have shown the presence of dipping reflectors, offsets, and other complexities. These studies have also shown that the upper crust is relatively transparent to vertical rays, whereas the lower crust presents multiple reflections (Fig. 11.13). This has been interpreted as a thin-layered or laminated structure in the form of multiple lenses of small dimensions in the lower crust. The upper crust is, then, formed by a more rigid or brittle material and the lower crust by one that is more ductile, laminated, and possibly with greater fluid concentrations. This result agrees with the hypothesis of a *seismogenic layer* (a layer where earthquakes are generated) that, more or less, coincides with the upper crust, as we will see in Chapter 21.

Table 11.1 Compositions of the most common igneous rocks.

	Granite (%)	Basalt (%)	Peridotite (%)
SiO ₂	72	51	43
Al ₂ O ₃	14	14	4
Fe _n O _m	3	12	12
MgO	1	6	34
CaO	2	10	4
Na ₂ O	3	2	1

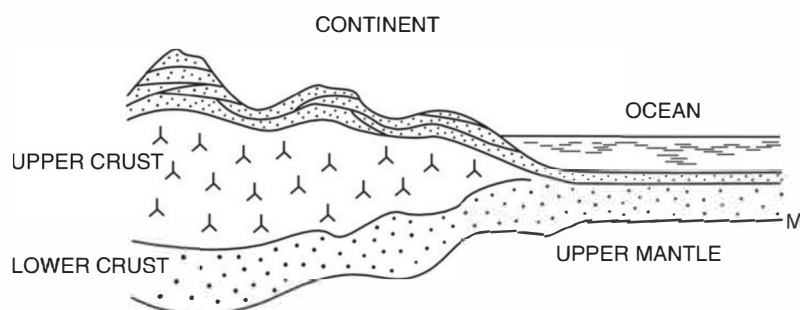


Fig. 11.12 Variation of the crust's thickness on going from oceanic to continental structure.

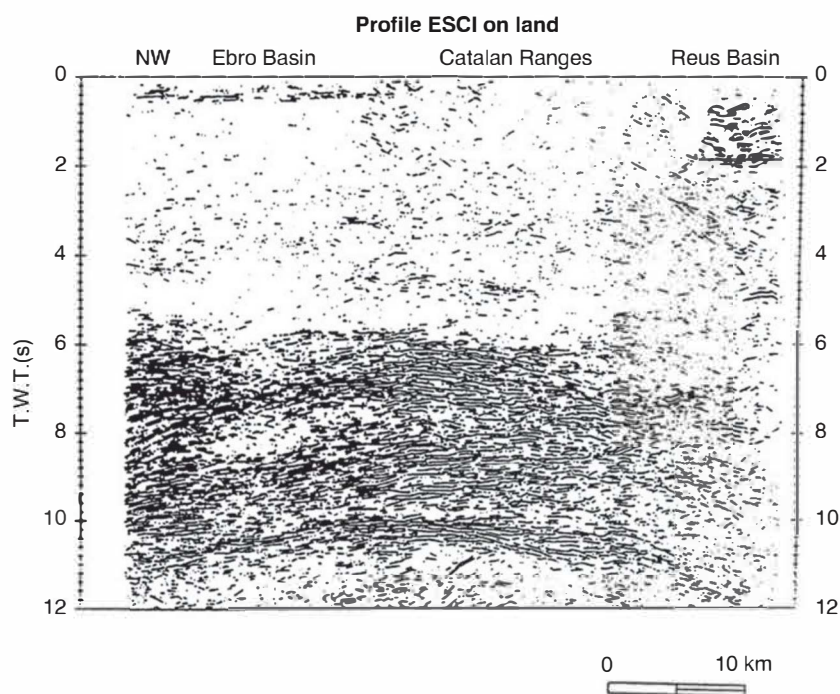


Fig. 11.13 A seismic record section for a vertical reflection profile in the Valencia trough (ESCI 1991, 1992). On the vertical scale are two-way times in seconds. The transparent upper crust and laminated lower crust are shown (Gallart *et al.*, 1995) (with permission from the Sociedad Geológica de España).

11.3.1 The mineralogical composition of the crust

Seismic observations only give us direct information on the distribution of velocity of body waves inside the Earth and indirectly the density and elastic coefficients. The mineralogical compositions of materials in the interior of the Earth can be derived from those of rocks on

the surface, laboratory experiments, and indirect evidence. Surface rocks can be separated according to their origin into sedimentary and igneous. Briefly, sedimentary rocks are formed by the compaction of sediments deposited mainly on the sea floor. The most abundant are limestones (CaCO_3) and sandstones (SiO_2). The seismic velocity in rocks varies according to the degree of compactness and depends on their age. Igneous rocks are formed by solidification from fused magma material. The most representative are granite, basalt, and peridotite. These three types of rocks are examples of acidic, basic or mafic, and ultrabasic or ultramafic rocks. This classification refers to the content of silica (SiO_2), which is greater than 52% for acid rocks, 45–52% for basic rocks, and less than 45% for ultrabasic rocks. Other classifications give for acid rocks a silica content of over 66% and call intermediate those with silica contents in the range 52–66%. The proportions of components of the three basic igneous rocks are given in [Table 11.1](#).

The most commonly accepted results give for the global composition of the continental crust a proportion of silica (SiO_2) of approximately 60%, with 64% in the upper crust and 58% in the lower crust. The content of Al_2O_3 is 16% and that of CaO is 7%. The contents of FeO and MgO are 9% and 5%, with greater proportions in the lower crust. The old denomination of granitic and basaltic layers for the upper and lower crust is, then, only an approximation to their mineralogical compositions. The composition of the oceanic crust is 49% SiO_2 , 16% Al_2O_3 , 11% CaO , 11% FeO , and 8% MgO . This composition is similar to that of the lower continental crust and to those of basaltic rocks.

11.4 The upper mantle

From the point of view of seismology, material under the crust down to a depth of 700 km forms the upper mantle ([Section 2.6](#)). The most important characteristics of the distribution of velocities of P and S waves in its interior are the following. Velocities increase with depth from values under the Moho of approximately $7.8\text{--}8.3\text{ km s}^{-1}$ for P and $3.8\text{--}4.1\text{ km s}^{-1}$ for S waves, to about 10.7 km s^{-1} for P and 5.9 km s^{-1} for S waves. Approximately between 60 and 220 km depth, a low-velocity layer, which is more pronounced for S waves, is found. The depth, thickness, and decrease in velocity of this layer vary from region to region and in some regions it may not even be present. This low-velocity layer is responsible for the shadow zone for P waves observed around 20° distance ([Section 9.6.3](#)). Under the low-velocity layer, from approximately 220 km depth, velocities both for P and S waves increase slowly. At depths of 450 and 670 km there are two zones of rapid increase in velocity ([Section 9.6.2](#)) ([Fig. 11.14a](#)). Travel time curves of rays that penetrate to those depths have properties corresponding to rapid increase in velocity with normal and retrograde branches ([Fig. 11.14b](#)). These two zones are sometimes considered to be discontinuities in the velocity distribution.

The material of the upper mantle, according to Alfred E. Ringwood, is a composite of magnesium and iron silicates, called *pyrolite*, formed from basalt and olivine (Ringwood, 1975). With increasing depth, there are changes of phase in the minerals, for example, that of plagioclase to form garnet. Changes in mineral phase with the same composition

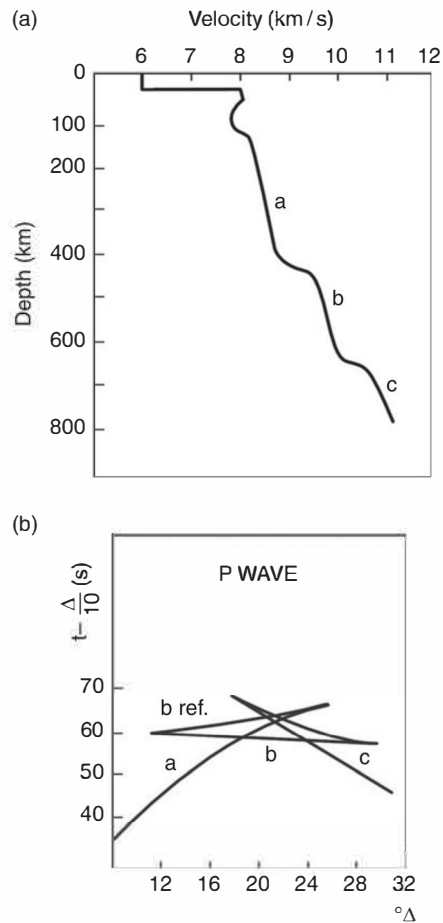


Fig. 11.14 (a) The velocity distribution of P waves in the upper mantle. (b) The reduced travel time curve, showing branches of waves refracted in a, reflected in b, refracted in b, reflected in c, and refracted in c.

(magnesium iron silicate, $(\text{Mg,Fe})\text{SiO}_3$) are also used to interpret the rapid increase in velocity at the 450 and 670 km discontinuities; for the first, a change in structure from olivine to spinel has been proposed, and, for the second, a change from spinel to perovskite.

11.4.1 The lithosphere and the asthenosphere

As we saw in Section 2.7, the crust and upper mantle are the parts active in tectonic processes. For this reason, we have introduced according to plate tectonics the terms *lithosphere* for the rigid layer formed by the crust and part of the upper mantle that forms the units of plate motion, and *asthenosphere* for the weak layer with plastic or viscous characteristics underneath (Fig. 11.15). The existence within the Earth of an upper rigid layer (the lithosphere) divided into plates lying over a plastic or viscous layer (the

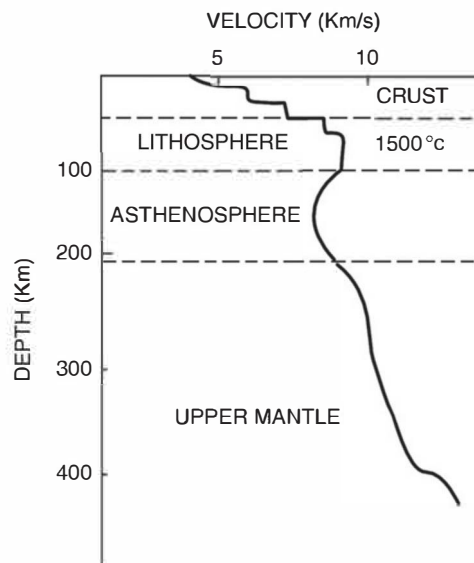


Fig. 11.15 P wave velocity profile in the lithosphere, asthenosphere, and mantle.

asthenosphere) that allows its horizontal motion is necessary in all models of plate tectonics (see further discussion in [Section 21.5](#)).

From the seismological point of view, the lithosphere can be identified with the layer that includes the crust and the lid of the upper mantle, where the velocities of P and S waves are relatively high. The thickness found for the lithosphere varies on going from oceans to continents, with values in the range 60–120 km. From the thermal point of view, the lithosphere is the relatively cold layer with a large temperature gradient and its lower boundary is located at the isothermal of 1500 °C. These three definitions of the lithosphere (tectonic, seismological, and thermal) do not always coincide and the term itself has no unique meaning. The asthenosphere is identified with the low-velocity layer for both P and S waves, which extends from depths of about 60 km under the oceans and 120 km under the continents to about 200 km down. Its thickness varies from region to region and its lower limit is not well defined.

The lower limit of the upper mantle at about 700 km depth coincides with the maximum depth of earthquakes and approximately with the 2000 °C isotherm. This depth corresponds also to the depth of material involved in tectonic processes and, for this reason, the whole layer down to this depth is sometimes called the *tectonosphere*. In relation to the upper mantle, subduction zones have a depth in some cases of as much as 700 km. These zones first manifested through the occurrence of deep earthquakes or *Wadati–Benioff zones*, where the oceanic lithosphere is introduced under the continental lithosphere in regions of collisions between plates ([Section 21.5](#)). Details of these and other heterogeneities in the upper mantle have recently been detected by seismic tomography. Inside the upper mantle, regions of positive (higher than normal velocities) and negative (lower than normal velocities) anomalies with respect to the models of radial symmetry, with

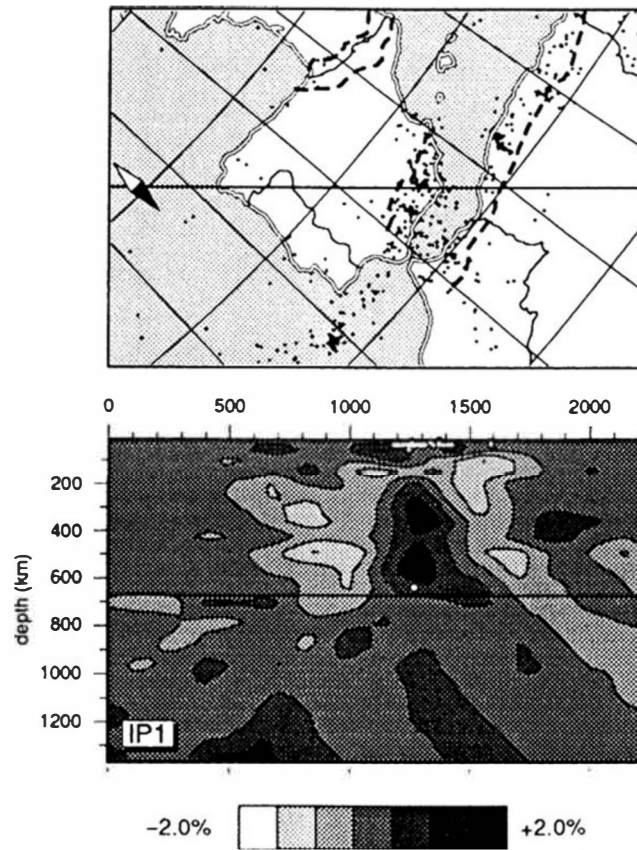


Fig. 11.16 A study of seismic tomography under the Iberian Peninsula. A positive-velocity anomalous region at depths between 200 and 600 km is shown (Blanco and Spakman, 1993) (with permission from Elsevier Science).

values up to 5%, have been found for P and S waves (Fig. 11.16). These regions are interpreted as zones in the upper mantle where material is colder (positive anomalies) and warmer (negative anomalies). Positive anomalies are associated with subduction zones and negative anomalies are associated with the material under rift or extension zones and are related to upward thermal convection currents (Woodward and Master, 1991). Researchers using seismic tomography have also discovered structures in the lower mantle, sometimes down to the core, that may show that tectonic processes are really more deeply rooted.

11.5 The lower mantle

In the lower mantle P wave velocities increase from 10.75 to 13.72 km s^{-1} and S wave velocities increase from 5.95 to 7.26 km s^{-1} . The density also increases in a gradual way

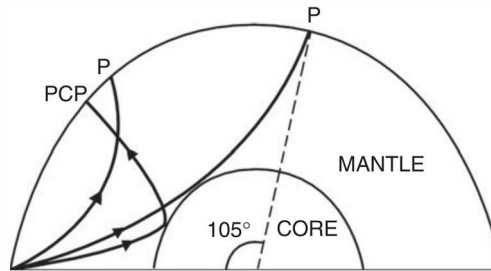


Fig. 11.17 Ray trajectories for P waves inside the mantle for direct rays and rays that are reflected at the core.

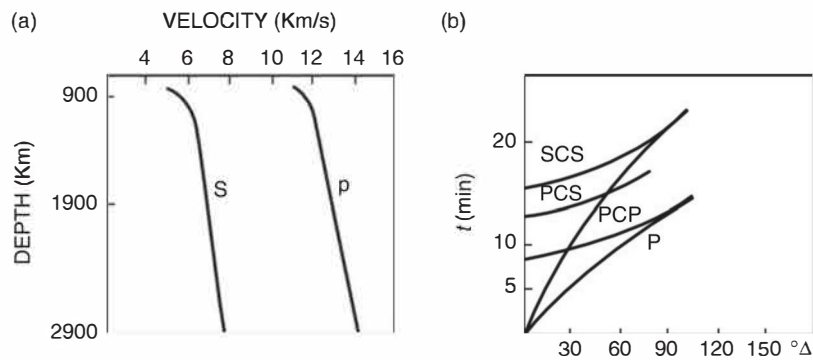


Fig. 11.18 (a) Velocity distributions of P and S waves in the lower mantle. (b) Travel time curves of P and S rays, direct and reflected at the core.

from 4.38 to 5.57 g cm^{-3} . The trajectories of seismic rays correspond to a normal distribution of velocities (Section 10.6.1). In the lower mantle we find direct P and S rays and rays that are reflected on the free surface and on the surface of the core.

Direct P and S rays propagated in the lower mantle arrive at distances between 35° and 105° (Fig. 11.17). At 105° there arrive rays that turn upward immediately above the core's boundary. Between 105° and 115° there is the shadow zone produced by the decrease of velocity in the core (Section 10.6.3), but we find arrivals of P and S waves that have been diffracted at the core's surface. Travel time curves for direct P and S waves have characteristics corresponding to a normal distribution of velocities such that the velocity increases slowly and its gradient changes very little (Section 10.6.1) (Fig. 11.18).

The strong velocity contrast at the CMB (transitions from 13.72 to 8.06 km s^{-1} for P waves, and from 8 km s^{-1} to zero for S waves) produces reflections of P and S waves that are called PcP and ScS waves (c for core) and also converted reflections, PcS and ScP (Fig. 11.18b). Travel time curves corresponding to reflected waves are concave upward and tangential to direct branches at their maximum distance (105°) (Fig. 11.18b).

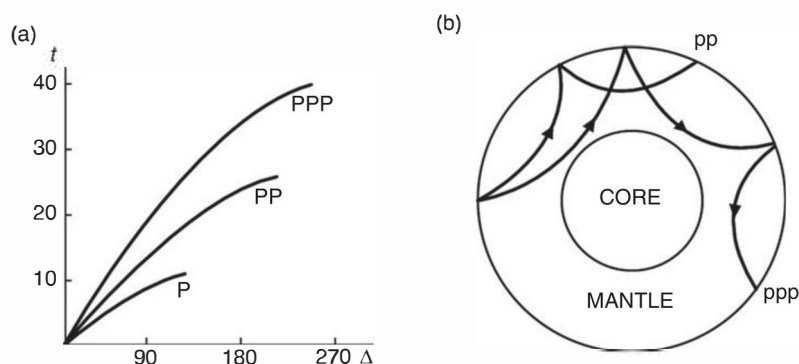


Fig. 11.19 (a) Ray trajectories for P waves reflected at the free surface. (b) Travel time curves.

For surface foci, intersection times for PcP and ScS are 8'30" and 15'36". Diffracted P and S waves arriving at distances greater than 105° can also be considered to be grazing reflected PcP and ScS waves at the core's surface.

Inside the lower mantle, we find also P and S waves that are reflected once or several times at the free surface (Fig. 11.19a). These waves are designated PP, PPP, etc. and SS, SSS, etc. When there are phase conversions in the reflections we have PS, SP, PSS, PSP, PPS, SSP, etc. Rays reflected several times arrive at later times than do direct rays and those reflected fewer times (Fig. 11.19b). One also observes waves reflected at the discontinuities in the upper mantle, that is, at depths of 260 (the lower limit of the low-velocity layer), 450, and 670 km. These reflected rays are designated PdP, where d is the depth of the reflecting layer (e.g. P450P). Owing to the small contrasts in velocity at these discontinuities, the energies of these waves are not large and they are difficult to detect on seismograms.

For deep-focus earthquakes, two types of rays reflected at the free surface can arrive at the same distance, corresponding to rays that travel upward or downward from the focus (Fig. 11.20a). In the first case, the rays are designated pP and they arrive at a given distance before the PP rays that correspond to rays that travel downward (Fig. 11.20b). The ray with take-off angle $i_h = \pi/2$ arrives at the minimum distance Δ_m (for example, for $h = 300$ km, $\Delta_m = 24^\circ$), where the two branches of travel times for pP and PP rays come together. Since the distance from the focus to the surface is small compared with the distance to the observation point, pP rays arrive a little later than do direct P rays. Hence the time interval $\Delta t = t(\text{pP}) - t(\text{P})$ depends on the focal depth and is used in its determination.

Rays that travel through the lower mantle can be observed on seismograms recorded at distances from 30° to 100°. For a surface focus, the most prominent body-wave phases correspond to direct P and S waves; the other phases that are observable are reflected on the surface and the core (Fig. 11.21). On seismograms of surface shocks the large amplitudes after S waves correspond to surface waves, as we will see in the next chapter (Fig. 11.21). On a seismogram of a deep-focus earthquake, we can also observe pP and sS phases and waves reflected at the core, PcP and ScS (Fig. 11.22). Usually, seismograms for deep-focus earthquakes provide clearer recordings of body waves because of the low generation of surface waves. Seismic tomography has been applied to the study of velocity heterogeneities in the

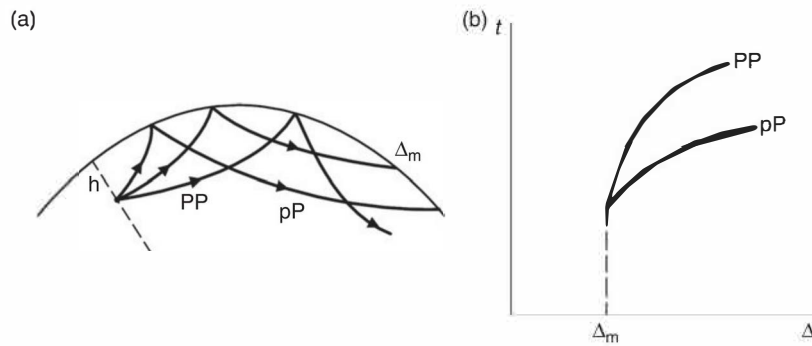


Fig. 11.20 (a) Ray trajectories of pP and PP waves. (b) Travel time curves.

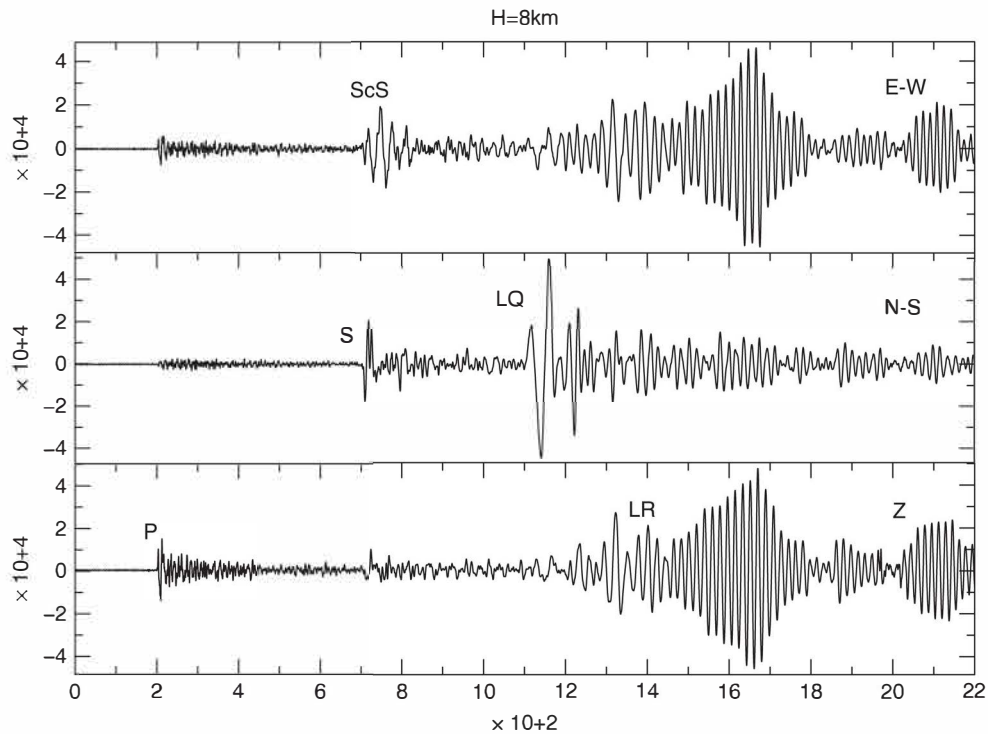


Fig. 11.21 A seismogram for a surface-focus earthquake in Haiti (12 January 2010) recorded at the IFR station ($\Delta = 62^\circ$), showing direct and reflected P and S phases and surface waves.

lower mantle (Inoue *et al.*, 1990, Dziewonski, 1996, Grant, 2002). These studies have found the presence of regions with higher and lower velocities than those of the reference model. As material is considered to be fairly homogeneous, high velocities (positive anomalies) are interpreted as corresponding to colder material and lower velocities (negative anomalies) to

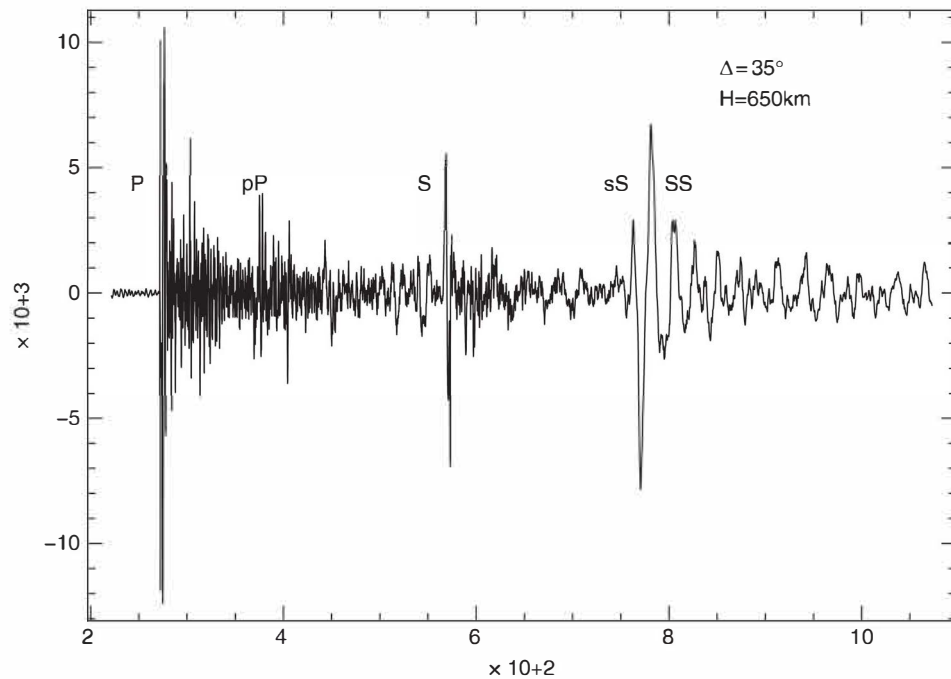


Fig. 11.22 A seismogram for a deep-focus earthquake in Granada (Spain), 11 April 2010, recorded at the KIV BB station, showing direct P and S phases, S phases reflected at the free surface.

warmer material. These results have been related to the presence in the lower mantle of convection currents with upward (warm) and downward (cold) movements (Fig. 11.23).

The composition of the lower mantle is considered to be very homogeneous and predominantly formed by silicates of magnesium and iron, and oxides of magnesium and iron with perovskite structure. The gradual increases in velocity and density inside the lower mantle are attributed to crystalline structures which become more compact with depth due to the increase in pressure. Inside the lower mantle the pressure increases from 10^4 to 10^5 MPa and temperatures increase from 1000 to 3000 °C.

11.6 The core

As we saw in Section 2.6 the core is formed by a fluid outer core and a solid inner core. Thus S waves do not propagate in the outer core, indicating that its material is fluid enough not to allow the existence of shear stress. The P wave velocity decreases sharply from 13.72 km s^{-1} at the base of the lower mantle to 8.06 km s^{-1} at the surface of the outer core (Fig. 11.24). According to Section 10.6.3, we expect to find a shadow zone in travel time curves and, after this, a duplication with two branches, one normal and another retrograde at a later time (Fig. 10.11e). For the Earth, without considering the inner core, the shadow zone extends from 105° to about 135° . Rays that penetrate into the core are called PKP rays (K from

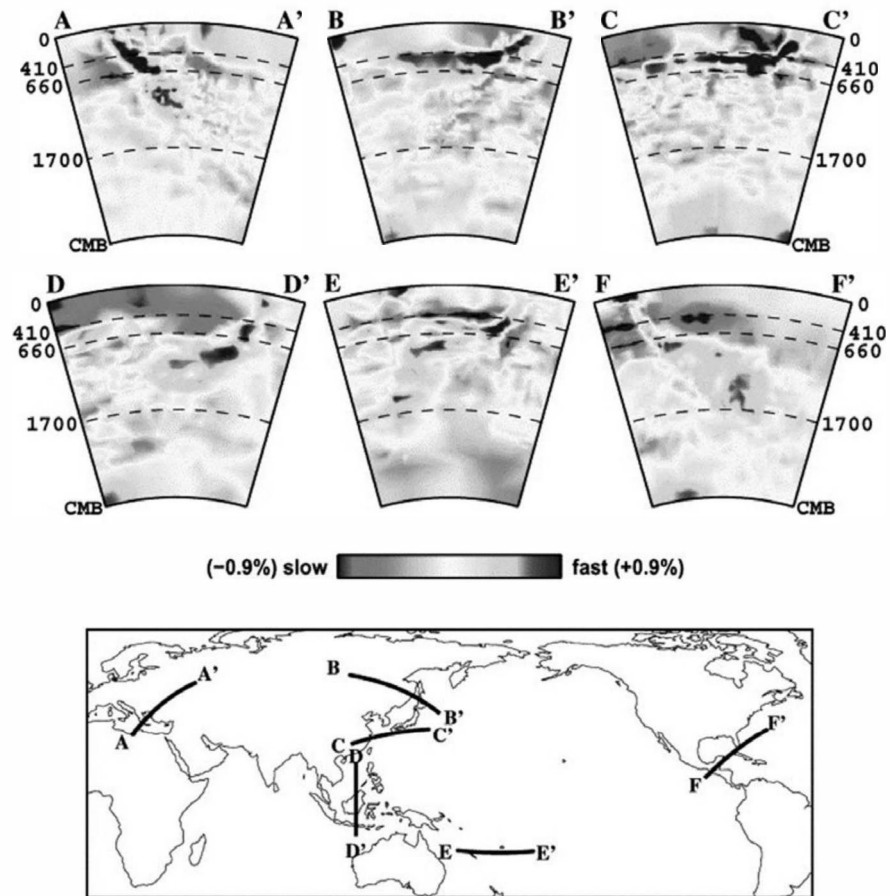


Fig. 11.23 Seismic tomography showing fast anomalies associated with subducted slabs (Romanowicz, 2003).

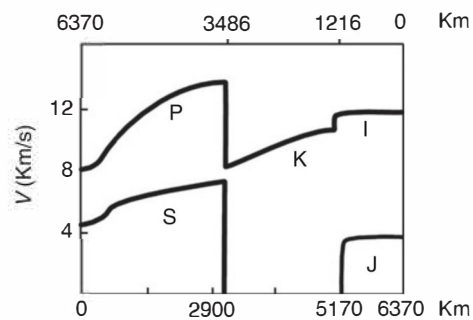


Fig. 11.24 Velocity distributions of P and S waves in the mantle, and in the outer and inner core.

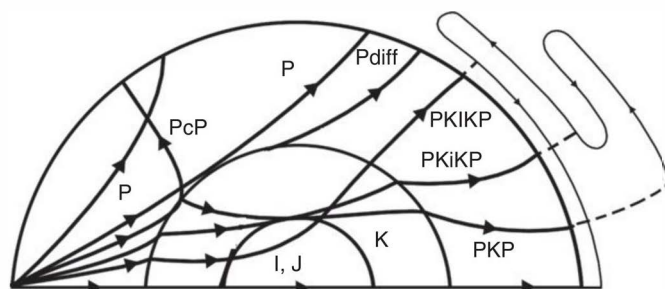


Fig. 11.25

Ray trajectories of P waves through the mantle, and of P waves refracted and reflected from the outer and inner core.

Kern, German for core). Owing to the rapid decrease in velocity, rays incident on the core with large angles of incidence are refracted to distances of about 170° . With decreasing angle of incidence, rays arrive at shorter distances, forming the retrograde branch PKP_2 , which is concave upward. The minimum distance for these rays is 135° . At this distance there starts also the normal branch PKP_1 that continues to 180° (Fig. 11.25). For a model with these characteristics, travel times in the absence of diffracted waves and inner-core reflections present a shadow zone from 105° to 135° . From 135° we find duplication of branches PKP_1 and PKP_2 with a time delay of 3 min with respect to direct P waves propagated in the lower mantle (Fig. 11.26).

The presence of the solid inner core with a greater P wave velocity than that of the outer core (Fig. 11.25) modifies the ray trajectories and travel time curves. First of all, there are rays reflected on the surface of the inner core (designated $PKiKP$) that start to arrive at a distance of about 120° , in the shadow zone (Fig. 11.26a). Arrivals of these waves form the main evidence for the existence of the inner core. Rays refracted in the inner core, $PKIKP$, arrive ahead of the PKP_1 rays. The travel time curve is, then, formed by four branches, PKP_1 , PKP_2 , $PKiKP$, and $PKIKP$ (Fig. 11.26b). PKP_1 rays arrive from 135° to 156° , PKP_2 (concave upward) rays from 135° to 170° , $PKiKP$ (concave upward) rays from 120° to 156° , and $PKIKP$ rays from 120° to 180° (Fig. 11.26b). The identification of these phases is difficult, for this reason it is convenient to use a stack of seismograms recorded at similar teleseismic distances. Figure 11.27 shows an example of the identification of $PKiKP$ and $PKIKP$ using this technique.

As has been mentioned, the fluid nature of the outer core blocks the propagation of S waves. However, S waves propagated in the lower mantle are converted into P waves in the outer core that are again converted into S waves in the mantle. Such rays, when they travel through the outer core only, are designated SKS rays; if they penetrate the inner core, they are called SKIKS rays. SKiKS rays are those reflected on the surface of the inner core (Fig. 11.28). Since, for rays passing from S (the lower mantle) to K (the outer core) and from K to I (the inner core), the velocity always increases, the travel times have characteristics associated with media with zones of rapid increase in velocity (Section 10.6.2, Fig. 10.10). Direct S rays arrive at distances of up to 105° . Rays refracted in the outer core, SKS rays, start to arrive at 62° , a point common to the branch of reflected ScS rays, and they are observed up to 133° . Rays that are refracted into the inner core, SKIKS rays, arrive at distances from 99° to

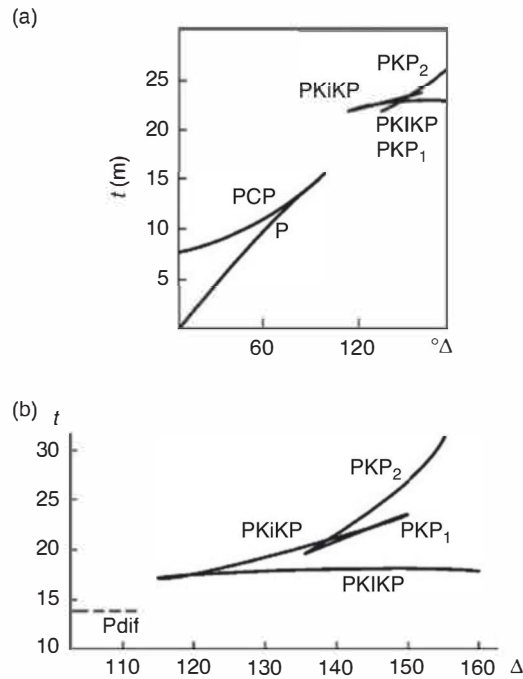


Fig. 11.26

(a) Traveling time curves for P waves in the mantle and core. (b) Detail of traveling times for P waves in the outer and inner core.

180°. Rays reflected from the inner core, SKiKS rays, arrive at distances from 99° to 133°, forming a concave upward branch (Fig. 11.29). Since the inner core is solid, S waves also propagate in its interior designated by the letter J. S waves in the mantle are converted into P (K) waves in the outer core and into S (J) waves in the inner core. They are called SKJKS rays, and their travel time branch is parallel to the SKIKS branch, but delayed in time. These waves have little energy and are difficult to observe.

On seismograms, for epicentral distances greater than 105°, we observe arrivals of rays that have traveled through the core. Owing to the complexity of arrivals of rays refracted and reflected in the outer and inner core, seismograms vary rapidly in appearance on going from one distance to another. PKiKP rays start to arrive at about 120°, PKIKP rays at 120°, and PKP₁ and PKP₂ rays at 135° (Fig. 11.30).

Velocity models of the core agree in their general characteristics with those shown in Fig. 11.31 (Bolt, 1982; Jeanloz, 1990; Souriau, 2009). The velocity of P waves decreases sharply from 13.5 km s⁻¹ at the base of the lower mantle to 8.06 km s⁻¹ in the outer core, where it increases slowly from 8.06 to 10.36 km s⁻¹, while the velocity of S waves is zero. There are several models for the boundary between the outer and inner core. A simple model consists in a small decrease in the velocity gradient followed by a sharp increase in velocity from 10.36 to 11.03 km s⁻¹. In the inner core, the velocity is practically constant or increases very slowly from 11.03 to 11.26 km s⁻¹. The S velocity in the inner core is about 4 km s⁻¹.

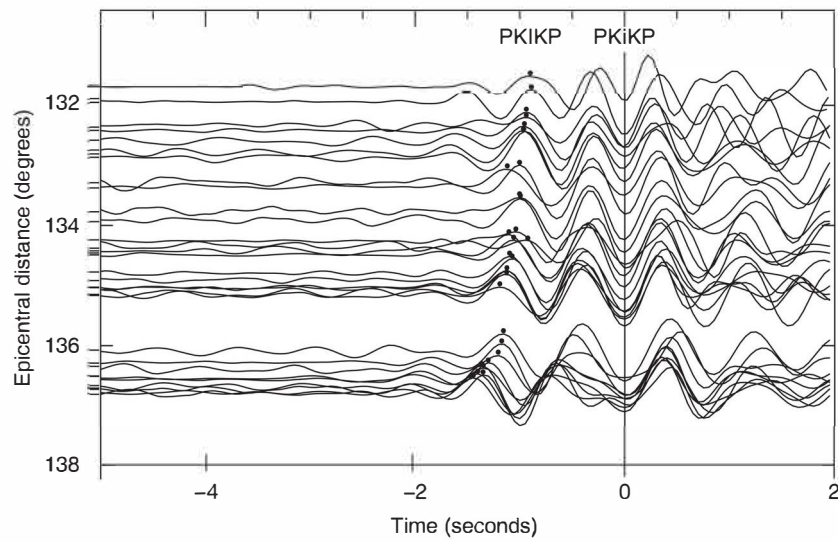


Fig. 11.27 Stack of seismograms of reflected and refracted phases on the inner core for the Colombia 26 April 1999 earthquake ($h = 164$ km). Black line PKIKP and black circles PKiKP (courtesy of M. Ramírez).

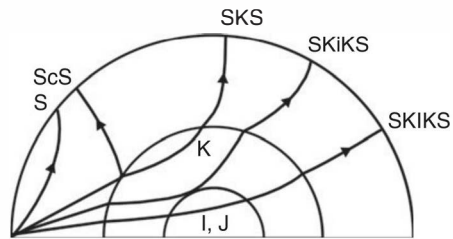


Fig. 11.28 Ray trajectories of S waves in the mantle and of S waves refracted in and reflected from the core.

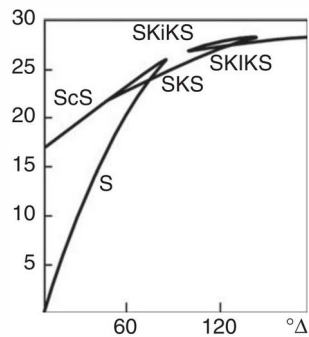


Fig. 11.29 Travel times for S waves in the mantle and for S waves refracted in and reflected from the outer and inner core.

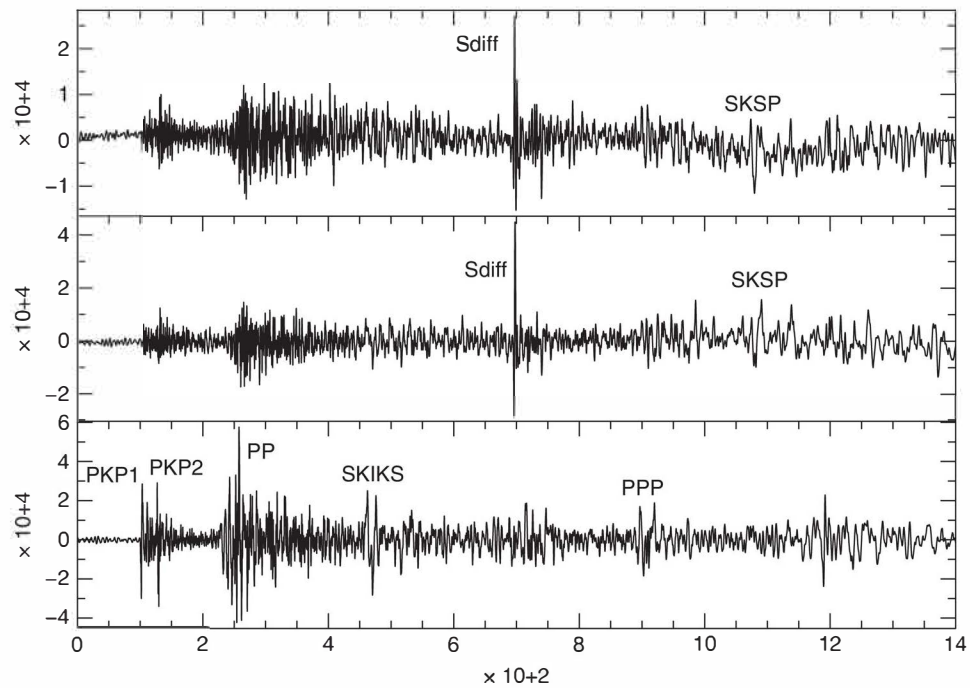


Fig. 11.30

A seismogram of a deep earthquake ($h = 552$ km) in Flores Sea (27 February 2015), recorded at the LPAZ station ($\Delta = 155^\circ$) showing the phases of the core. (Top EW, middle NS, bottom Z.)

At the CMB, density increases rapidly from 5.57 g cm^{-3} at the base of the lower mantle to 9.9 g cm^{-3} in the outer core. Inside the core, the density increases gradually, with a small increment at the boundary of the inner core, reaching a value of 13.1 g cm^{-3} . Owing to its large density, the mass of the core is approximately 35% of the mass of the Earth. The composition of the core is basically iron with a small proportion (about 10%) of lighter elements such as silicon, nickel, sulfur, carbon, and oxygen. The presence of some of these elements is necessary in order to explain the values of the density and bulk modulus in the core, both of which are lower than expected for pure iron at high pressure and temperature. The iron composition is confirmed by the presence of this mineral in meteorites and the existence of the Earth's internal magnetic field. Generation of this field is explained by models of auto-excited dynamos produced by the motion of fluid conductive material in the outer core. The material of the core seems to be very homogeneous and the difference between the outer and inner core is due only to its fluid or solid state. Some authors, however, also propose a difference in composition, with iron and nickel in the inner core and iron and sulfur in the outer core.

Transition zones between the mantle and the outer core, CMB, called by Bullen the *D zone*, and between the outer and inner cores, or the *F zone*, are still a subject for discussion. First models of the CMB posited a sharp decrease in velocity, whereas more recent models favor a previous increase in velocity at a depth of about 2600 km followed by the decrease (Young and Lay, 1987). The CMB marks a sharp contrast in physical and chemical characteristics, from the solid material of magnesium and iron silicates in the

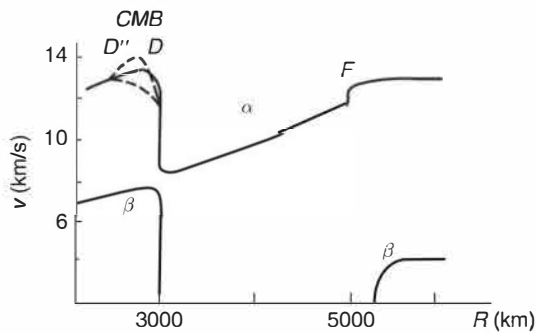


Fig. 11.31 Velocity distributions of P and S waves in the inner and outer core. Transition zones between the core and the mantle (D and D'') (the CMB) and between the inner and the outer core (F) are shown.

mantle to a fluid made basically of iron in the outer core. Above this boundary, there is an anomalous layer in the mantle known as the D'' zone, up to 300 km in thickness and with great lateral inhomogeneities. The characteristics of the outer–inner core boundary or F zone are less well known. Models present a gradual increase in velocity from about 4600 km depth (Fig. 11.31). The transition zone may extend further, from 4500 to 5200 km depth (Song and Helmberger, 1992). Modern seismic tomography studies have revealed the existence of lateral velocity anomalies that may be related to thermal convection currents present in the fluid material in the outer core. These currents are related to generation of the magnetic field. Velocity anomalies and a strong axial anisotropy have also been found in the solid inner core (Chapter 15).

11.7 Summary

Methods based on the analysis of seismic waves have been applied to determine the structure of the Earth's interior. From these, global models of the distribution of P and S wave velocities have been determined, such as JB and PREM models. For the structure of crust and upper mantle refraction, wide-angle reflection, vertical reflection, and seismic tomography have been used. The crust, divided into upper and lower, has different structures for oceanic, continental, and orogenic regions, with thickness ranging from 10 to 60 km and P velocities about 6.6 km/s and S velocities about 4 km/s. The structure of the mantle and the core is derived from the analysis of travel time and amplitude of waves generated by earthquakes. The upper mantle extends to 700 km depth with two regions of rapid increase of velocity at 450 km and 670 km. The lower mantle extends from 700 km to 2900 km depth, where is located the boundary of the core. Its composition is very uniform with a gradual increase in P (8 to 13.7 km/s) and S wave velocities (4.4 to 7.3 km/s).

The core extends from 2900 km depth to the center. It is divided into outer and inner cores. The outer core is liquid and no S waves propagate. P waves have a sharp decrease in velocity (from 13.7 to 8 km/s). P and S waves are reflected at the core boundary (PcP and

ScS) and P waves are transmitted through the outer core (PKP) and the inner core (PKIKP). S waves incident on the core boundary are transmitted as P waves through the outer core and then converted again to S waves in the mantle (SKS). Through the inner core there are both P and S waves transmitted (SKIKS and SKJKS). Modern tomographic studies show a three-dimensional structure in the mantle and core.

11.8 Problems

- P11.1 In the seismograms of EM 11_1, identify the arrival time of the main phases.
- P11.2 Consider the Earth's crust, composed of two layers of thickness 12 and 18 km, and constant P wave velocities of 6 and 7.5 km/s, respectively, on top of a semi-infinite mantle of constant velocity 8 km/s. There is a seismic focus in the upper crust at a depth of h below the surface. Using the seismogram of EM Fig._P11.2, recorded at a station at 150 km of epicentral distance, identify the arrivals of P_n and P_g and calculate the depth of the focus.
- P11.3 The seismogram of EM Fig._P11.3 corresponds to a deep earthquake. The P wave velocity at the focus is 12.4 km/s and Poisson's ratio is 0.25. Identify the P, pP, and sP and calculate the focal depth.
- P11.4 Consider a spherical Earth of radius $R = 3000$ km and constant P wave speed of propagation of 6 km/s. Within it there is a core of radius $R/2$ and speed of propagation $v = \frac{a}{\sqrt{r}}$. The seismogram of EM Fig._P11.4 has been recorded at a station of epicentral distance Δ and it corresponds to an earthquake with focus at depth $R/6$ from the Earth's surface. Given that the Poisson ratio is $1/4$, calculate the epicentral distance assuming the same path for P and S waves.

The presence of a free surface on an elastic medium, as is the case on Earth, introduces a series of phenomena that must be considered in the study of waves produced by earthquakes, as observed in seismograms. First of all, as we have seen in [Section 6.4](#), on a free surface there are body-wave reflections. Under certain conditions (supercritical incidence of S waves) the generation of inhomogeneous or evanescent P waves occurs ([Sections 6.3](#) and [6.4](#)). These are body waves that propagate along a direction parallel to the free surface and whose amplitudes decrease with the distance from the free surface. A different phenomenon, which we will consider now, is the generation of surface waves from constructive interference of body waves in connection with a free surface. Surface waves are defined as those produced in media with a free surface which propagate parallel to the surface and whose amplitudes decrease with the distance from the surface. They are generated by energy brought to the free surface by incident body waves. Their existence is related to the presence of a free surface, and other surfaces of contact between layers of different elastic properties. They are very important in seismology as they produce very large amplitudes of surface ground motion. There are also waves of similar characteristics, but not related to a free surface, which are associated with an interface between two media in contact.

12.1 Rayleigh waves in a half-space

The first problem is that of determining whether, in an elastic, homogeneous half-space limited by a plane $x_3 = 0$ (x_3 being positive upward), there exist surface waves, that is, waves that propagate in the direction of x_1 with velocity c and whose amplitudes decrease with depth ($-x_3$) ([Fig. 12.1](#)). According to [\(5.89\)](#) and [\(5.90\)](#), the components of displacement u_1 and u_3 can be expressed in terms of the scalar potentials ϕ and ψ , and u_2 is kept apart:

$$u_1 = \phi_{,1} - \psi_{,3}, \quad (12.1)$$

$$u_2 = u_2, \quad (12.2)$$

$$u_3 = \phi_{,3} + \psi_{,1}. \quad (12.3)$$

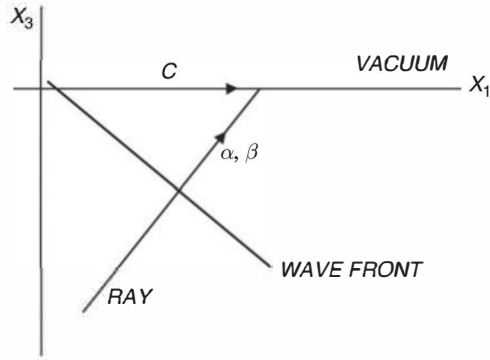


Fig. 12.1 A ray incident on a free surface and the component of velocity along the surface.

For waves of frequency ω that propagate in the positive x_1 direction with velocity c , the potentials, ϕ and ψ , and transversal displacement, u_2 , are given by

$$\phi = f(x_3) \exp[ik(x_1 - ct)], \quad (12.4)$$

$$\psi = g(x_3) \exp[ik(x_1 - ct)], \quad (12.5)$$

$$u_2 = h(x_3) \exp[ik(x_1 - ct)], \quad (12.6)$$

where $f(x_3)$, $g(x_3)$, and $h(x_3)$ express the amplitude's dependence on depth and $k = \omega/c$ is the wave number. By substitution into the wave equations for ϕ , ψ , and u_2 (5.58), we obtain

$$f''(x_3) + k^2 r^2 f(x_3) = 0, \quad (12.7)$$

$$g''(x_3) + k^2 s^2 g(x_3) = 0, \quad (12.8)$$

$$h''(x_3) + k^2 s^2 h(x_3) = 0, \quad (12.9)$$

where, just like in (6.60) and (6.61),

$$r = \left(\frac{c^2}{\alpha^2} - 1 \right)^{1/2}, \quad (12.10)$$

$$s = \left(\frac{c^2}{\beta^2} - 1 \right)^{1/2}. \quad (12.11)$$

The solutions of equations (12.7)–(12.9) are

$$f(x_3) = A' \exp(ikrx_3) + A \exp(-ikrx_3), \quad (12.12)$$

and expressions of the same form for $g(x_3)$ and $h(x_3)$, with r replaced by s .

Surface waves must have amplitude that decreases with depth ($-x_3$). Hence r and s must be imaginary and positive and consequently, $A' = B' = C' = 0$.

The problem is related to the incidence of waves on a free surface. According to (12.10) and (12.11), if r and s are imaginary, $c < \beta < \alpha$, that is, the velocity of surface waves is smaller than that of S waves. This is the opposite to what occurs for incident body waves on the free surface, where c is the apparent velocity in the direction of x_1 , r and s are real, and $c > \alpha > \beta$ (Section 6.4). In this case, $r = \tan e$ and $s = \tan f$ (6.60 and 6.61), where e and f are the angles of emergence of incident and reflected P and S waves. If $\alpha > c > \beta$, there are incident and reflected S waves and inhomogeneous P waves (r is imaginary). As we have seen, for surface waves, we must have both r and s imaginary and $c < \beta < \alpha$. By substituting $f(x_3)$ from (12.12) and similar expressions for $g(x_3)$ and $h(x_3)$ into (12.4)–(12.6), with the conditions specified above, we finally obtain

$$\phi = A \exp[-ikrx_3 + ik(x_1 - ct)], \quad (12.13)$$

$$\psi = B \exp[-iksx_3 + ik(x_1 - ct)], \quad (12.14)$$

$$u_2 = C \exp[-iksx_3 + ik(x_1 - ct)]. \quad (12.15)$$

To evaluate the arbitrary constants A , B , and C , we apply the boundary conditions at the free surface, that is, null stress across the surface: for $x_3 = 0$, $\tau_{31} = \tau_{32} = \tau_{33} = 0$. For an isotropic medium, the stresses in terms of the displacements, according to (4.20), are

$$\tau_{31} = \mu(u_{3,1} + u_{1,3}), \quad (12.16)$$

$$\tau_{32} = \mu(u_{3,2} + u_{2,3}), \quad (12.17)$$

$$\tau_{33} = (\lambda + 2\mu)u_{3,3} + \lambda(u_{1,1} + u_{2,2}). \quad (12.18)$$

By replacing the displacements u_1 and u_3 in terms of the potentials (12.1) and (12.3), and considering that they are independent of x_2 , for $x_3 = 0$, we obtain the following equations:

$$2\phi_{,31} + \psi_{,11} - \psi_{,33} = 0, \quad (12.19)$$

$$u_{2,3} = 0, \quad (12.20)$$

$$(\lambda + 2\mu)\phi_{,33} + \lambda\phi_{,11} + 2\mu\psi_{,13} = 0. \quad (12.21)$$

By substitution of (12.15) into equation (12.20), we obtain that $C = 0$. Therefore, surface waves in a half-space do not have a horizontal transverse component of displacement (u_2). By substituting (12.13) and (12.14) into (12.19) and (12.21), we obtain for $x_3 = 0$ that

$$2rA - (1 - s^2)B = 0, \quad (12.22)$$

$$[\alpha^2(r^2 + 1) - 2\beta^2]A - 2\beta^2sB = 0. \quad (12.23)$$

This is a homogeneous system of equations; therefore, the condition for the existence of a solution, apart from the trivial $A = B = 0$, is that the determinant of the system be null:

$$\begin{vmatrix} 2r & -(1 - s^2) \\ \alpha^2(r^2 + 1) - 2\beta^2 & -2\beta^2s \end{vmatrix} = 0. \quad (12.24)$$

Solving the determinant gives

$$[\alpha^2(r^2 + 1) - 2\beta^2](1 - s^2) - 4rs\beta^2 = 0. \quad (12.25)$$

On replacing the values of r and s from (12.10) and (12.11), and taking into account that they are imaginary, we have

$$\left(2 - \frac{c^2}{\beta^2}\right)^2 = 4\left(1 - \frac{c^2}{\alpha^2}\right)^{1/2} \left(1 - \frac{c^2}{\beta^2}\right)^{1/2}. \quad (12.26)$$

This equation is known as *Rayleigh's equation*, in honor of John William Strutt, Lord Rayleigh, who solved this problem for the first time in 1887. In order to study the solutions of Rayleigh's equation, we make changes of variables, $\xi = (c/\beta)^2$ and $q = (\beta/\alpha)^2$. The value of q is related to Poisson's ratio (Section 4.2) and depends on the elastic properties of the medium. On taking the square of (12.26) and eliminating the solution $\xi = 0$, we obtain a cubic equation for ξ :

$$\xi^3 - 8\xi^2 + 8(3 + 2q)\xi + 16(q - 1) = 0. \quad (12.27)$$

Solutions of this equation depend on the value of q . Since Poisson's ratio has values between 0 and $\frac{1}{2}$ (Section 4.2), q varies between $\frac{1}{2}$ and 0. For each value of q , equation (12.27) has three roots. From these only those for values of $\xi < 1$ correspond to surface waves, since they must satisfy the condition that $c < \beta$. Waves corresponding to these solutions are called *Rayleigh waves* and their velocity $c_R = \sqrt{\xi} \beta$ is a fraction of the shear-wave velocity in the half-space. In Table 12.1, values of ξ and c_R corresponding to certain values of σ are given, together with the relations among the values of q , λ , and μ .

Values of ξ and q versus σ are shown in Fig. 12.2. For all possible values of σ , ξ varies very little, between 0.7640 and 0.9128. For $\sigma = \frac{1}{2}$, the medium is a liquid, μ and β are null, and the root $\xi = 0.9128$ corresponds to a zero value of c_R . Hence, Rayleigh waves do not exist in a liquid half-space.

A special case is that for $\sigma = \frac{1}{4}$ ($\lambda = \mu$) and $q = \frac{1}{3}$, a condition which is approximately satisfied for the Earth's materials. For this case, we derived in Section 6.4 expressions for the reflection coefficients of incident P and S waves reflected on the free surface of an elastic medium, V_{PP} (6.100) and V_{SS} (6.106). Both (6.100) and (6.106) have the same numerator. If we put this numerator equal to zero, we obtain the same equation (12.26) for $q = \frac{1}{3}$. This shows again that the generation of Rayleigh waves is related to the reflection of

Table 12.1 Values of ξ and c_R corresponding to certain values of σ , together with the relations among the values of q , λ , and μ .

σ	q	λ	μ	ξ	c_R
0	$\frac{1}{2}$	0	$\frac{3}{2} K$	0.7640	0.8741β
$\frac{1}{8}$	$\frac{2}{7}$	$\frac{1}{3} \mu$	K	0.8059	0.8977β
$\frac{1}{4}$	$\frac{1}{3}$	μ	$\frac{3}{5} K$	0.8453	0.9194β
$\frac{1}{2}$	0	K	0	0.9128	0.0

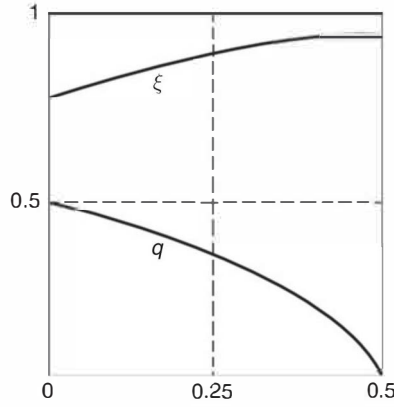


Fig. 12.2 Variation of $q(\beta/a)$ and $\xi(c/\beta)$ with Poisson's ratio.

P and S waves on the free surface of an elastic half-space. For this particular case, equation (12.27) results in

$$\xi^3 - 8\xi^2 + \frac{56}{3}\xi - \frac{32}{3} = 0. \quad (12.28)$$

The roots of this equation are 4, $2 + 2/\sqrt{3}$, and $2 - 2/\sqrt{3}$. The first two ($\xi > 1$) correspond to cases of wave reflection for which $c > \alpha > \beta$. For incident P waves, these values of ξ correspond to angles of incidence for which there are no reflected P waves, since $V_{PP} = 0$. Because $\cos e = \alpha/c$, these angles are $e = 30^\circ$ and $e = 12^\circ 47'$. For an incident S wave, they correspond to $f = 60^\circ$ and $f = 55^\circ 44'$ ($V_{SS} = 0$). The third root is $\xi = 0.8453$ and corresponds to Rayleigh waves of velocity 0.9194β .

In conclusion, in an elastic half-space, Rayleigh waves are generated that propagate parallel to the free surface with a velocity $c_R = \sqrt{\xi}\beta$, where ξ are the roots with values less than unity of equation (12.27) for given values of β/α . The amplitudes of these waves decrease exponentially with depth. In a certain way, these waves may be considered to be generated by energy brought to the surface by incident P and S waves that produce no reflections.

12.1.1 Displacements of Rayleigh waves

Displacements of Rayleigh waves are obtained by substituting expressions (12.13) and (12.14) for the potentials ϕ and ψ into (12.1) and (12.3). For the particular case of $\sigma = \frac{1}{4}$, we have $c_R = 0.9194\beta$, $r = 0.85i$, and $s = 0.39i$. By first substituting these values into (12.22), we obtain $B = -1.47iA$. Finally, taking only the real part, the displacements are given by

$$u_1 = -Ak(e^{0.85kx_3} - 0.58e^{0.39kx_3}) \sin[k(x_1 - c_R t)], \quad (12.29)$$

$$u_3 = -Ak(-0.85e^{0.85kx_3} + 1.47e^{0.39kx_3}) \cos[k(x_1 - c_R t)]. \quad (12.30)$$

We must remember that A is the potential amplitude (in units of m^2) and Ak is the displacement amplitude (in units of meters). For points at the free surface ($x_3 = 0$), letting $a = -Ak$, we find that

$$u_1 = 0.42a \sin[k(x_1 - c_R t)], \quad (12.31)$$

$$u_3 = 0.62a \cos[k(x_1 - c_R t)]. \quad (12.32)$$

Since Rayleigh waves have no transverse component, they are polarized in the vertical plane. The horizontal and vertical components are shifted in phase by $\pi/2$ and thus the motion is elliptical. If we substitute into (12.31) and (12.32) the values of t during a complete cycle (0 to T , where $T = 2\pi/\omega$ is the period), we obtain for the particle's motion an ellipse with a vertical major axis and retrograde motion (opposite to that of wave propagation) (Fig. 12.3).

The dependence of the displacement components u_1 and u_3 on depth ($-x_3$) is given by (12.29) and (12.30). There is a value of x_3 for which u_1 is null, $x_3 = -0.19\lambda$ ($\lambda = 2\pi/k$ is the wave length), whereas u_3 is never null. At the depth where u_1 is null, its amplitude changes

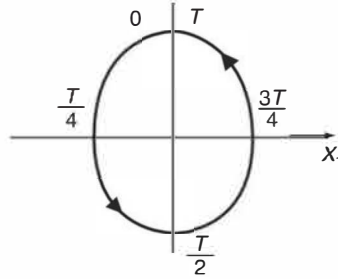


Fig. 12.3 Diagram of a particle's motion on the vertical plane for a Rayleigh wave at the free surface.

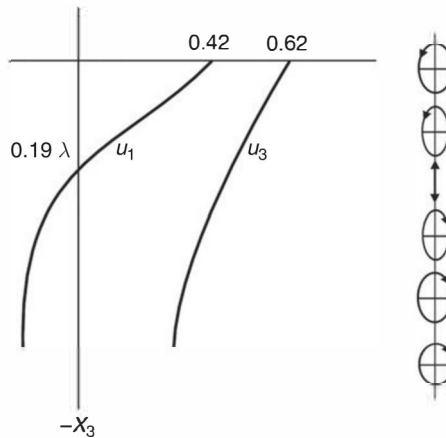


Fig. 12.4 Plot of the amplitudes of displacement components u_1 and u_3 of Rayleigh waves in a half-space, versus depth.

sign. For greater depths the particle's motion is prograde (Fig. 12.4). With increasing depth, the amplitudes of u_1 and u_3 decrease exponentially, with u_3 always larger than u_1 .

12.2 A liquid layer over a rigid half-space. Guided waves

The existence of one or more layers of finite thickness and different properties lying over a half-space modifies the results we have obtained for Rayleigh waves in a half-space. As an introduction to this problem, we consider a liquid layer of thickness H , density ρ , and velocity α , lying over a rigid half-space (Fig. 12.5). Since there is no propagation of waves in a rigid medium, we have no surface waves in the half-space, only waves contained in the liquid layer.

The motion in the liquid layer can be expressed in terms of a scalar potential ϕ . For waves that propagate in the positive x_1 direction with velocity c , the potential is given by

$$\phi = (A e^{ikrx_3} + B e^{-ikrx_3}) e^{ik(x_1 - ct)}, \quad (12.33)$$

where r is given by (12.10). Since x_3 can vary only between 0 and H , we cannot yet impose any condition on the real or imaginary character of r . The boundary condition on the free surface ($x_3 = H$) is that the normal component of stress across the surface be null, $\tau_{33} = 0$. At the base of the layer ($x_3 = 0$), the vertical component of the displacement is null, $u_3 = 0$. In terms of the potential ϕ , these two conditions are

$$x_3 = H : \quad \phi_{,11} + \phi_{,33} = 0 \quad (12.34)$$

$$x_3 = 0 : \quad \phi_{,3} = 0. \quad (12.35)$$

The first condition, from the relation among the cubic dilation θ , the Laplacian of ϕ (4.86), and the wave equation (5.3), can be written as $\rho\omega^2\phi = 0$. Using the potential ϕ according to (12.3), conditions (12.34) and (12.35) give

$$Ae^{ikrH} + Be^{-ikrH} = 0, \quad (12.36)$$

$$A - B = 0. \quad (12.37)$$

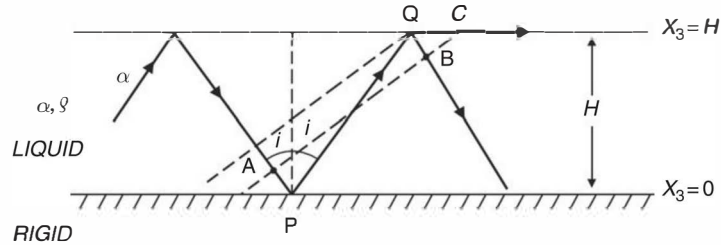


Fig. 12.5

Rays in a liquid layer over a rigid medium, showing constructive interference on the wave front AB.

Since this is a homogeneous system of equations for A and B , the existence of a solution implies that its determinant is null:

$$e^{ikrH} + e^{-ikrH} = 0. \quad (12.38)$$

If r is imaginary, equation (12.38) leads to the impossible result, $\cosh(krH) = 0$. Therefore, r must be real and $c > \alpha$, and condition (12.38) gives

$$\cos(krH) = 0. \quad (12.39)$$

On substituting for r its value in (12.10), the condition leads to

$$kH(c^2/\alpha^2 - 1)^{1/2} = \left(n + \frac{1}{2}\right)\pi, \quad n = 0, 1, 2, 3, \dots \quad (12.40)$$

Since r is real, that is $c > \alpha$, waves in the x_1 direction with velocity c are formed by reflections in the interior of the liquid layer. Waves propagate parallel to both limits of the layer and are called *guided* or *channeled waves*. Contrary to the previous case (12.26), in equation (12.40) the wave number k appears. This implies that the velocity of propagation is a function of the wave number, $c(k)$, or of the frequency, $c(\omega)$. This means that waves are dispersed, that is, waves of different frequency travel with difference velocities and arrive at different times producing a train of waves. Equation (12.40), which relates the velocity to the frequency, is called the *dispersion equation* (Sections 5.2.2 and 5.4.1). Wave dispersion is a consequence of introducing a finite dimension (layer thickness) into the problem, since the velocity depends on the relation between the wave length and the layer's thickness.

Another important result is the presence in the dispersion equation (12.40) of the integer n which takes an infinite number of discrete values. For each value of n , the relation between c and k is different; that is, we have a different type of wave propagation. Each of these forms of wave propagation is called a *mode*. Here the presence of modes is due to the finite dimension of the layer's thickness. The lowest value of n corresponds to the *fundamental mode* and the other values to *higher modes* or *harmonics*. The same phenomenon appears in the problem of vibrations of finite elastic bodies (Chapter 14).

12.2.1 Constructive interference

We have seen that the solution for guided waves corresponds to real values of r , that is, $c > \alpha$. Therefore we can consider that these waves are generated by *constructive interference* of body (acoustic) waves propagated in the interior of the liquid layer and reflected from its two surfaces. Hence we can also find the dispersion equation (12.40) by using the ray theory approach for plane waves and the condition of constructive interference. This condition implies that waves that coincide on the same wave front must be in phase so that their amplitudes are summed. For a ray that is reflected on both surfaces (Fig. 12.5), the same wave front AB corresponds to rays that pass through A and through B. For constructive interference on this wave front, the two waves must be in phase; that is, the distance along the ray from A to B ($AP + PQ + QB$) must be an

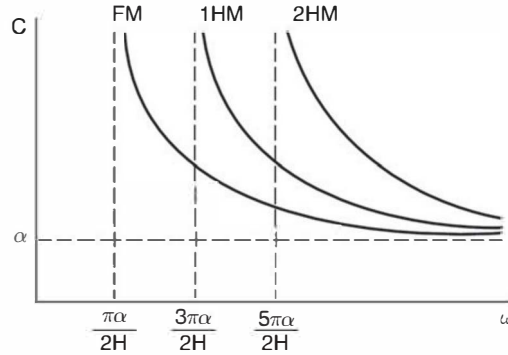


Fig. 12.6

Dispersion curves for guided waves in a liquid layer over a rigid medium. The fundamental mode and the two first higher modes with their cut-off frequencies are shown.

integer multiple of the wave length, taking into account the possible phase shifts at the points of reflection. The phase shift at the rigid surface is zero whereas that at the free surface of the liquid is π (Section 6.2.5). The condition of constructive interference can, then, be written as

$$\frac{2\pi}{\lambda_\alpha} (AP + PQ + QB) - \pi = 2\pi n. \quad (12.41)$$

According to the ray geometry (Fig. 12.5),

$$AP + PQ + QB = 2H \cos i. \quad (12.42)$$

Since $\sin i = \alpha/c$, $\cos i = (\alpha/c)(c^2/\alpha^2 - 1)^{1/2}$, and $k = k_\alpha \sin i$, by substitution into (12.41) we obtain the dispersion equation (12.39). Guided waves that appear on the liquid layer are, thus, formed by constructive interference of acoustic waves reflected from its two surfaces. In an elastic layer, for example in the Earth's continental crust, there are similar channelled S waves, called Lg waves, that are prominent in short distance seismograms (Fig. 11.10).

12.2.2 The dispersion equation and curves

If we write (12.40) in terms of the frequency ω and solve for c , we find the explicit form of the dispersion equation:

$$c(\omega) = \frac{1}{\left\{ \frac{1}{\alpha^2} - \left[\left(n + \frac{1}{2} \right) \pi \right]^2 \frac{1}{H^2 \omega^2} \right\}^{1/2}}. \quad (12.43)$$

The fundamental mode corresponds to $n = 0$; thus, according to (12.40), $kHr = \pi/2$. Higher modes correspond to $n \geq 1$. For real values of $c(\omega)$, the expression in the denominator must be larger than zero. The frequency ω_c corresponding to the zero value of the denominator is

called the *cut-off frequency*, since there are no values of $c(\omega)$ for $\omega < \omega_c$. For any mode of order n , the cut-off frequency is

$$\omega_c = \frac{\pi(n + \frac{1}{2})\alpha}{H}. \quad (12.44)$$

The lower cut-off frequency, $\omega_c = \alpha\pi/(2H)$, corresponds to the fundamental mode. In all modes, the velocity c becomes infinite for the cut-off frequency $\omega = \omega_c$. With increasing ω , the velocity c decreases and, in the limit, when ω tends to infinity, c tends to α . The $c(\omega)$ curves are called *dispersion curves* (Fig. 12.6). For each mode there is a dispersion curve with values of the velocity for frequencies from the cut-off frequency to infinity. We can see that the lowest value of the velocity corresponds to the fundamental mode for a given frequency and increases with the order of the mode. Also, a given velocity corresponds to the lowest frequency for the fundamental mode and to higher frequencies with increasing order of the modes (Fig. 12.6).

12.2.3 Displacements

Using equations (12.33) and (12.37), we can determine the displacements of guided waves, by taking derivatives of the potential ϕ , according to (12.1) and (12.3):

$$u_1 = -2Ak \cos\left(krH \frac{x_3}{H}\right) \sin[k(x_1 - ct)], \quad (12.45)$$

$$u_3 = -2Akr \sin\left(krH \frac{x_3}{H}\right) \cos[k(x_1 - ct)]. \quad (12.46)$$

The components u_1 and u_3 are shifted in phase by $\pi/2$, and the resulting motion in the vertical plane is elliptical and retrograde. For the fundamental mode, $kHr = \pi/2$, we obtain

$$u_1 = -2AK \cos\left(\frac{\pi x_3}{2H}\right) \sin[k(x_1 - ct)], \quad (12.47)$$

$$u_3 = -\frac{\pi A}{H} \sin\left(\frac{\pi x_3}{2H}\right) \cos[k(x_1 - ct)]. \quad (12.48)$$

For the free surface ($x_3 = H$), $u_1 = 0$ and the motion is vertical, whereas at the base of the layer ($x_3 = 0$), $u_3 = 0$ and the motion is horizontal. Inside the layer, the motion is elliptical; the major axis changes with depth from vertical to horizontal (the motion is circular for $x_3 = (2H/\pi) \tan^{-1} [\pi/(2hk)]$) (Fig. 12.7).

For the fundamental mode, there is no depth inside the layer at which the amplitude either of u_1 or of u_3 is zero (Fig. 12.7a). For higher modes, there are values of x_3 inside the layer, equal to the order number of the mode, at which either u_1 or u_3 is zero (nodes), but they do not coincide. For example, for the first higher mode, $krH = 3\pi/2$, u_1 is zero for $x_3 = H/3$ and u_3 is zero for $x_3 = 2H/3$ (Fig. 12.8).

In conclusion, guided waves in a liquid layer over a rigid half-space present the following characteristics that are found in all cases of the propagation of guided waves in layers of finite thickness: (a) dispersion; that is, the velocity depends on the frequency; (b) an infinite

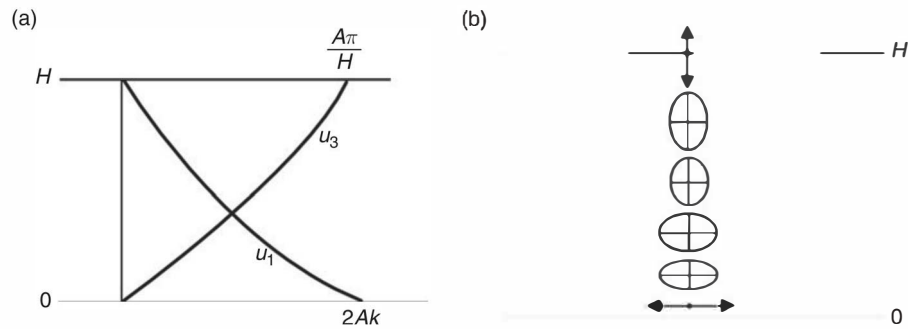


Fig. 12.7

(a) The amplitude of the displacement components u_1 and u_3 of guided waves in a liquid layer over a rigid medium for the fundamental mode. (b) Diagram showing a particle's motion on the vertical plane.

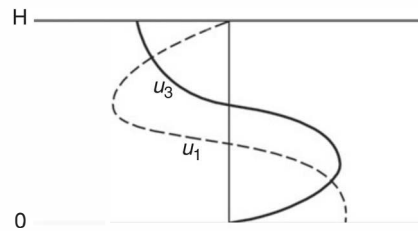


Fig. 12.8

Amplitudes with depth for the displacement components u_1 and u_3 for the first higher mode of guided waves in a liquid layer over a rigid medium.

number of modes corresponding to values of n , namely fundamental ($n = 0$) and higher ($n \geq 1$) modes; (c) for each mode there is a different dispersion curve with a cut-off frequency; and (d) displacements for higher modes have a number of amplitude nodes equal to their order number.

12.3 An elastic layer over a half-space. Love waves

The characteristics of Rayleigh waves on the free surface of an elastic half-space do not agree with those of surface waves observed on the Earth. Ever since the installation of the first seismographs, it has been observed that surface waves are formed by dispersed trains and have transverse horizontal components of motion. Therefore, the theory deduced in [Section 12.1](#) is not sufficient to explain observations and we must introduce the presence of layers that affect the propagation of surface waves. This is due in the Earth to the presence of the crust and other layers ([Chapter 11](#)). The first to consider this problem was Love in 1911. He found that, in an elastic layer over a half-space, dispersed surface waves are produced with horizontal transverse components, which are now called *Love waves*.

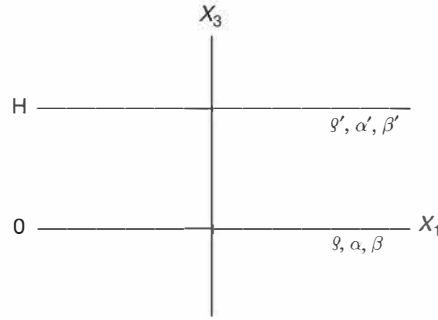


Fig. 12.9 An elastic layer of thickness H over an elastic half-space.

In this problem, we consider only waves with a transverse component of displacement, u_2 , that propagate in the direction x_1 with a velocity c , in a medium consisting of a layer of thickness H , density ρ' , and shear velocity β' , over a half-space of density ρ and velocity β (Fig. 12.9). We can write the displacements in the form of (12.15), but allowing for waves traveling in positive and negative directions of x_3 inside the layer leads to

$$u'_2 = (A' e^{iks'x_3} + B' e^{-iks'x_3}) e^{ik(x_1 - ct)}, \quad (12.49)$$

$$u_2 = B e^{-iksx_3 + ik(x_1 - ct)}, \quad (12.50)$$

where s is given by (12.11) and s' has a similar form obtained by replacing β' by β . In (12.50), according to the condition for surface waves, amplitudes must decrease with depth ($-x_3$); therefore, s must be positive and imaginary, $s = i\bar{s}$ ($c < \beta$). We cannot impose any condition on s' , since x_3 inside the layer is limited to values between 0 and H . The boundary conditions are now as follows: (a) at the free surface ($x_3 = H$), the component of stress τ'_{32} is null; (b) at the surface of contact between the layer and the half-space ($x_3 = 0$), there is continuity of the stress ($\tau'_{32} = \tau_{32}$) and displacement ($u'_2 = u_2$). The second condition implies that the layer and half-space are welded together. The three boundary conditions in terms of the displacements are

$$x_3 = H : \mu' u'_{2,3} = 0, \quad (12.51)$$

$$x_3 = 0 : \mu' u'_{2,3} = \mu u_{2,3}, \quad (12.52)$$

$$u'_2 = u_2. \quad (12.53)$$

By substitution into (12.49) and (12.50), we find the following three equations:

$$A' e^{iks'H} - B' e^{-iks'H} = 0, \quad (12.54)$$

$$A' \mu' s' - B' \mu' s' + B s \mu = 0, \quad (12.55)$$

$$A' + B' - B = 0. \quad (12.56)$$

A solution of this system of equations, apart from the trivial $A' = B' = B = 0$, requires that the determinant be zero:

$$\begin{vmatrix} e^{iks'H} & -e^{-iks'H} & 0 \\ \mu's' & -\mu's' & \mu s \\ 1 & 1 & -1 \end{vmatrix} = 0. \quad (12.57)$$

By expanding the determinant we obtain

$$\frac{\mu s}{\mu's'} = \frac{e^{iks'H} - e^{-iks'H}}{e^{iks'H} + e^{-iks'H}}. \quad (12.58)$$

If s' is imaginary ($c < \beta'$), then, on making the substitution $s' = i\bar{s}'$, we obtain

$$\frac{\mu\bar{s}'}{\mu's'} = -\tanh(k\bar{s}'H). \quad (12.59)$$

Since the left-hand side is always positive and the hyperbolic tangent is too, this result is not possible. Therefore s' must be real ($\beta > c > \beta'$), so the velocity of the layer must be less than that of the half-space, and the velocities of Love waves have values between the two. Since s is imaginary and s' is real, from equation (12.58) we obtain

$$\frac{\mu(1 - c^2/\beta'^2)^{1/2}}{\mu'(c^2/\beta'^2 - 1)^{1/2}} = \tan[kH(c^2/\beta'^2 - 1)^{1/2}]. \quad (12.60)$$

On making the substitution $k = \omega/c$, we obtain

$$\frac{\mu(1/c^2 - 1/\beta'^2)^{1/2}}{\mu'(1/\beta'^2 - 1/c^2)^{1/2}} = \tan[\omega H(1/\beta'^2 - 1/c^2)^{1/2}]. \quad (12.61)$$

Just as in the case of guided waves in a liquid layer, the velocity of Love waves is a function of the frequency, $c(k)$, or $c(\omega)$, and equations (12.60) and (12.61) are dispersion equations. The tangent has positive values between zero and infinity for various intervals of its argument kHs' : the first between 0 and $\pi/2$, the second between π and $3\pi/2$, etc. Each of them corresponds to a mode of propagation, as we saw with guided waves; the fundamental mode corresponds to the first interval ($0 \leq khs' \leq \pi/2$) and higher modes to the rest.

12.3.1 Constructive interference

In a similar manner to what we did in the case of guided waves in a liquid layer, we can also deduce the dispersion equation of Love waves (12.60), using the principle of constructive interference for SH rays reflected inside the layer corresponding to angles of incidence greater than the critical one. The condition of constructive interference requires that waves on the same wave front must be in phase. In Fig. 12.10, at points A and B, waves must have the same phase; that is, the length $AP + PQ + QB$ must be a multiple of the wave length,

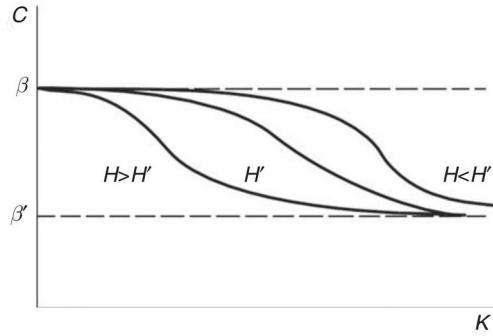


Fig. 12.11 Dispersion curves for the fundamental mode of Love waves for various thicknesses (H) of the layer.

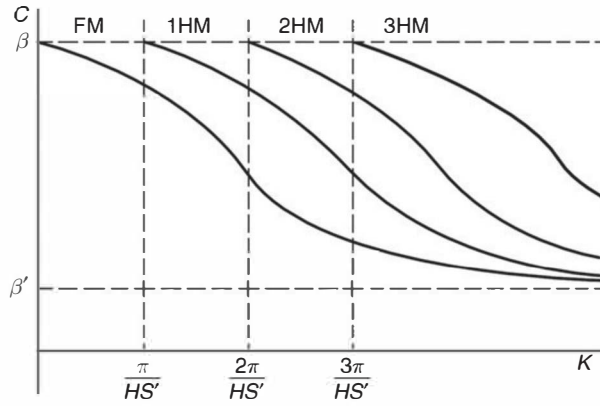


Fig. 12.12 Dispersion curves of Love waves in a layer over a half-space for the fundamental mode and the three first higher modes with their cut-off frequencies.

For the fundamental mode, according to (12.60), we have

$$0 \leq kH(c^2/\beta'^2 - 1)^{1/2} \leq \pi/2.$$

For $kHs' = 0$, $k = 0$ and $c = \beta$. For $kHs' = \pi/2$, $k = \infty$ and $c = \beta'$. In the fundamental mode all frequencies are present ($0 \leq k \leq \infty$) and the velocity of Love waves varies between the values of the S wave velocity in the half-space and the layer ($\beta' \leq c \leq \beta$). Thus, waves with lower frequencies (longer periods) arrive earlier than those of higher frequencies (shorter periods). This fact is used to identify the beginning of Love (LQ) waves and also of Rayleigh (LR), as we will see in the [next section](#), in a seismogram (Fig. 11.21). For the same velocities, the form of the curve depends on the layer thickness H (Fig. 12.11). The influence of the layer thickness (H) is conditioned by its relation to the wave length (λ). For $\lambda < H$, propagation is mainly influenced by the layer and $c \simeq \beta'$. For $\lambda > H$, propagation is conditioned by the half-space and $c \simeq \beta$. Therefore, for small thicknesses, the Love wave velocity approaches that of the half-space, $c \simeq \beta$; for relatively high

frequencies ($\omega > 2\pi c/H$), and for large thicknesses this is true for lower frequencies (Fig. 12.11).

For higher modes, since the argument of the tangent does not start at zero, there is a limit for low frequencies, called the cut-off frequency (Section 12.2). For each higher mode, the cut-off frequency has a different value. For a mode of order n , the cut-off wave number is

$$k_n = \frac{n\pi}{Hs'} \quad (12.63)$$

For the first higher mode we have

$$\pi \leq kH(c^2/\beta'^2 - 1)^{1/2} \leq 3\pi/2.$$

For $kHs' = \pi$, $k = \pi/(Hs')$, and $c = \beta$. For $kHs' = 3\pi/2$, $k = \infty$, and $c = \beta'$. The cut-off frequency is $k_1 = \pi/(Hs')$ and the form of the dispersion curve is similar to that of the fundamental mode (Fig. 12.12).

In general, dispersion curves for higher modes start at the cut-off frequency with the value $c = \beta$, and, for very high frequencies, this tends toward $c = \beta'$. For a given frequency, the lowest velocity corresponds to the fundamental mode and the highest to the mode of highest order. A given velocity corresponds to lower frequencies in the fundamental mode and to higher frequencies in higher modes (Fig. 12.12). For a given distance, waves of lower frequencies of the fundamental mode arrive at the same time as those of higher frequencies of higher modes.

12.3.3 Displacements

Displacements of Love waves can be deduced from expressions (12.49) and (12.50). From (12.54), we have

$$B' = A' e^{i2ks'H} \quad (12.64)$$

On substituting this into (12.49) and multiplying and dividing by $\exp(ikHs')$, we obtain for the real part of the displacement in the layer,

$$u'_2 = 2A' \cos\left[ks'H\left(1 - \frac{x_3}{H}\right)\right] \cos[k(s'H + x_1 - ct)]. \quad (12.65)$$

Displacements in the half-space can be derived from (12.50). From (12.56) and (12.64) we have

$$B = 2A' \cos(ks'H) e^{iks'H} \quad (12.66)$$

By substitution into (12.50), we obtain for the real part

$$u_2 = 2A' \cos(ks'H) e^{k\bar{x}x_3} \cos[k(s'H + x_1 - ct)]. \quad (12.67)$$

In the half-space, displacements decrease exponentially with depth ($-x_3$) (12.67). Inside the layer, displacements vary with depth depending on the value of $ks'H$ (12.65). Therefore, the variation will be different for each mode.

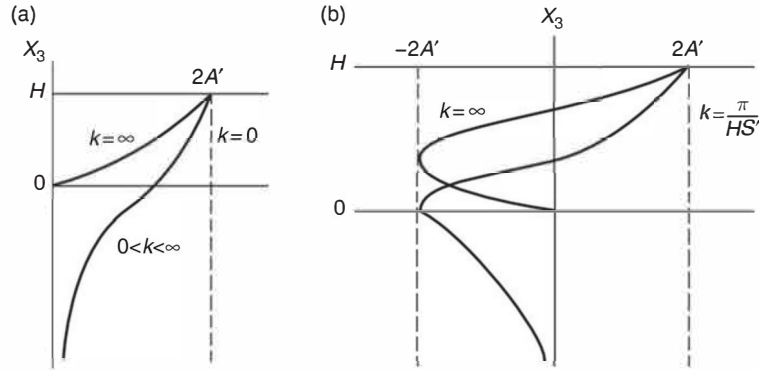


Fig. 12.13

The amplitude versus depth of Love waves in a layer over a half-space, for (a) the fundamental mode and (b) the first higher mode.

For the fundamental mode, $0 \leq kHs' \leq \pi/2$:

$$\begin{aligned} k = 0, \quad |u'_2| &= 2A' \text{ for all values of } x_3 \\ k = \infty, \quad |u'_2| &= \begin{cases} 2A' & \text{for } x_3 = H \\ 0 & \text{for } x_3 = 0 \end{cases} \end{aligned}$$

On the free surface ($x_3 = H$), $|u'_2| = 2A'$ for all values of k . At the base of the layer ($x_3 = 0$), the amplitude varies between $2A'$ and zero ($2A' > |u'_2| > 0$), for wave numbers between zero and infinity ($0 < k < \infty$). There are no nodes or values of x_3 inside the layer for which the amplitude is null (Fig. 12.13a). In the half-space, amplitudes decrease exponentially with depth, starting with the value at the base of the layer.

For the first higher mode, $\pi \leq kHs' \leq 3\pi/2$:

$$\begin{aligned} k = \pi/Hs', \quad |u'_2| &= \begin{cases} 2A' & \text{for } x_3 = H \\ 0 & \text{for } x_3 = H/2 \\ -2A' & \text{for } x_3 = 0 \end{cases} \\ k = \infty, \quad |u'_2| &= \begin{cases} 2A' & \text{for } x_3 = H \\ 0 & \text{for } x_3 = 2H/3 \\ -2A' & \text{for } x_3 = H/3 \\ 0 & \text{for } x_3 = 0 \end{cases} \end{aligned}$$

At the free surface, amplitude is always $|u'_2| = 2A'$, for all frequencies and modes. Depending on the value of k , there is a node where amplitudes are zero, for x_3 between $H/2$ and $3H/2$ (Fig. 12.13b). At the base of the layer, amplitudes have values between 0 and $-2A'$ for k between π/Hs' and infinity.

In a similar way, we can find the positions of nodes for any higher mode. In each mode, the number of nodes is equal to the order number of the mode. The dispersion and existence of modes of propagation in Love waves, just like in the case of guided waves in a liquid layer (Section 12.2), are consequences of the finite dimension of the layer thickness. The characteristics of propagation depend on the relation between the wave length and the layer thickness.

12.4 An elastic layer over a half-space. Rayleigh waves

In the [previous section](#) (Love waves), we considered only the transverse component of displacements (u_2); now we must consider the radial horizontal and vertical components (u_1 and u_3). The resulting surface waves are similar to the Rayleigh waves found for an elastic half-space ([Section 12.1](#)). To determine the existence and properties of these waves we proceed in the same way as for the problem of Love waves. The displacement components u_1 and u_3 are derived from the scalar potentials ϕ and ψ according to (12.1) and (12.3). For surface waves that propagate in the x_1 direction with a velocity c , in a layer of thickness H , and velocities α' and β' , over a half-space of velocities α and β ([Fig. 12.9](#)), taking into account that amplitudes must decrease with depth in the half-space, the potentials are given by

$$\phi' = A' \exp[ikr'x_3 + ik(x_1 - ct)] + B' \exp[-ikr'x_3 + ik(x_1 - ct)], \quad (12.68)$$

$$\psi' = C' \exp[iks'x_3 + ik(x_1 - ct)] + D' \exp[-iks'x_3 + ik(x_1 - ct)], \quad (12.69)$$

$$\phi = A \exp[-ikrx_3 + ik(x_1 - ct)], \quad (12.70)$$

$$\psi = C \exp[-iksx_3 + ik(x_1 - ct)], \quad (12.71)$$

where r and s are given by (12.10) and (12.11), and similarly r' and s' (on replacing α by α' and β by β'). Since amplitudes must decrease with depth ($-x_3$), r and s must be imaginary and positive ($r = i\bar{r}$ and $s = i\bar{s}$), and consequently $c < \beta < \alpha$.

Just as in the case of Love waves, the boundary conditions are $x_3 = H$, for the free surface, the components of stress are null, $\tau'_{31} = \tau'_{33} = 0$, the contact surface is $x_3 = 0$, and one has continuity of the displacement and stress components $u'_1 = u_1$, $u'_3 = u_3$, $\tau'_{31} = \tau_{31}$, and $\tau'_{33} = \tau_{33}$. As functions of the displacements, the boundary conditions result in the following equations:

$$x_3 = H :$$

$$\tau'_{31} = 0 \rightarrow u'_{3,1} + u'_{1,3} = 0, \quad (12.72)$$

$$\tau'_{33} = 0 \rightarrow \lambda' u'_{1,1} + (\lambda' + 2\mu') u'_{3,3} = 0. \quad (12.73)$$

$$x_3 = 0 :$$

$$u'_1 = u_1, \quad (12.74)$$

$$u'_3 = u_3, \quad (12.75)$$

$$\mu(u'_{3,1} + u'_{1,3}) = \mu(u_{3,1} + u_{1,3}), \quad (12.76)$$

$$\lambda u'_{1,1} + (\lambda + 2\mu) u'_{3,3} = \lambda u_{1,1} + (\lambda + \mu) u_{3,3}. \quad (12.77)$$

On substituting for the displacements in terms of the potentials ϕ , ϕ' , ψ , and ψ' , according to equations (12.68)–(12.71), we obtain

$$2r'(A'e^{ikr'H} - B'e^{-ikr'H}) + (1 - s'^2)(C'e^{iks'H} + D'e^{-iks'H}) = 0, \quad (12.78)$$

$$[\lambda'(1 + r'^2) + 2\mu'r'^2](A'e^{ikr'H} + B'e^{-ikr'H}) + 2\mu's'(C'e^{iks'H} - D'e^{-iks'H}) = 0, \quad (12.79)$$

$$A' + B' - s'C' + D's' = A + sC, \quad (12.80)$$

$$r'A' - r'B' + C' + D' = -rA + C, \quad (12.81)$$

$$\mu'[2r'(A' + B') + (1 - s'^2)(C' + D')] = \mu[2rA + (1 - s^2)C], \quad (12.82)$$

$$[\lambda'(1 + r'^2) + 2\mu'r'^2](A' + B') - 2\mu's'(C' - D') = [\lambda(1 + r^2) + 2\mu r^2]A + 2\mu sC. \quad (12.83)$$

Equations (12.78)–(12.83) form a system of six homogeneous equations for six unknowns A' , B' , C' , D' , A , and C , the amplitudes of the potentials ϕ and ψ in the layer and half-space. Again, the condition for a solution is that the determinant of the system is null. Making the determinant equal to zero, we find an equation for c , the velocity of Rayleigh waves. Since this equation implies that the velocity $c(k)$ is a function of the frequency, Rayleigh waves in a layer over a half-space are dispersed.

For the particular case when Poisson's ratio in the layer and half-space is $\sigma = \frac{1}{4}$ ($\lambda = \mu$ and $\lambda' = \mu'$), the expressions are simplified. We can verify that

$$\lambda(1 + r^2) + 2\mu r^2 = -\mu(1 + 3r^2) = \mu(1 - s^2). \quad (12.84)$$

If in the system of equations (12.78) to (12.83), we put $a' = kr'H$, and $b' = ks'H$, the determinant of the system can be written in the form

$$\begin{vmatrix} 2r'e^{ia'} & -2r'e^{-ia'} & (1 - s'^2)e^{ib'} & (1 - s'^2)e^{-ib'} & 0 & 0 \\ -(1 - s'^2)e^{ia'} & -(1 - s'^2)e^{-ia'} & 2s'e^{ib'} & 2s'e^{-ib'} & 0 & 0 \\ 1 & 1 & -s' & s' & -1 & -s \\ r' & -r' & 1 & 1 & r & -1 \\ 2\mu'r' & -2\mu'r' & \mu'(1 - s'^2) & \mu'(1 - s'^2) & -2\mu r & -\mu(1 - s^2) \\ -\mu'(1 - s'^2) & -\mu'(1 - s'^2) & 2\mu'r' & -2\mu'r' & \mu(1 - s^2) & -2\mu s \end{vmatrix} = 0.$$

The dispersion equation is obtained by expanding the determinant and putting it equal to zero. There are several ways of expanding this determinant. That proposed by Love in 1911 results in the equation

$$\xi'\eta - \xi\eta' = 0, \quad (12.85)$$

where, substituting the variables used here in place of those used by Love,

$$\xi = (1 - s'^2)[X \cos a' + (Yr/r') \sin a'] + 2s'[Wr \sin b' - Z/s' \sin b'], \quad (12.86)$$

$$\zeta' = (1 - s'^2)[Ws \cos \alpha' + Zr' \sin \alpha'] + 2s'[X \sin \alpha' - (Ys/s') \cos b'], \quad (12.87)$$

$$\eta = (1 - s'^2)[Wr \cos b' + (Z/s') \sin b'] + 2r'[X \sin \alpha' - (Yr/r') \cos \alpha'], \quad (12.88)$$

$$\eta' = (1 - s'^2)[X \cos b' + (Ys/s') \sin b'] + 2r'[Ws \sin \alpha' - (Z/r') \cos \alpha'], \quad (12.89)$$

and, using $g = \mu/\mu'$,

$$X = gc^2/\beta'^2 - 2(g - 1), \quad (12.90)$$

$$Y = c^2/\beta'^2 + 2(g - 1), \quad (12.91)$$

$$Z = gc^2/\beta'^2 - c^2/\beta'^2 - 2(g - 1), \quad (12.92)$$

$$W = 2(g - 1). \quad (12.93)$$

For a solution that implies the existence of Rayleigh waves that decrease exponentially with depth in the half-space, r and s are imaginary, while s' and r' are real. In some cases, however, r' can also be imaginary. These conditions imply that the velocity c of Rayleigh waves satisfies the condition $\alpha > \beta > c > \beta'$, while α' is, generally, less than c , but not necessarily. In a similar way to Love waves, Rayleigh waves in a layer over a half-space are formed by constructive interference of both P and SV waves reflected super-critically from the contact surface.

Solutions of the dispersion equation (12.85) give the velocity of Rayleigh waves as a function of the frequency $c(k)$ or $c(\omega)$, instead of a constant value (Section 12.1). There are an infinite number of solutions corresponding to different modes. Modes in this case are separated into two types, symmetric (M_1) and antisymmetric (M_2), in analogy with the vibrations of an elastic layer. For symmetric modes, vertical displacements at the free surface and contact surface have opposite signs, whereas for antisymmetric modes they have the same sign. The particle motion is elliptical with a vertical major axis. At the free surface, the particle motion is retrograde for symmetric modes and prograde for antisymmetric modes.

The characteristics of dispersion curves are different for each mode and depend on the values of the model parameters H, β, β', α , and α' . The first two symmetric modes are M_{11} and M_{12} and the first two antisymmetric ones are M_{21} and M_{22} (Fig. 12.14). M_{11} represents the fundamental mode, which is a symmetric mode and has all values of frequency ($0 \leq k \leq \infty$). All other modes, symmetric and antisymmetric, have a cut-off frequency (k_{1n}, k_{2n}). For M_{11} in the limit of low frequencies ($k = 0$) the velocity tends toward the Rayleigh wave velocity in the half-space ($c_R = 0.92\beta$), whereas for all the other modes the velocity for cut-off frequencies tends to that of S waves in the half-space (β). In the limit of high frequencies ($k = \infty$) for M_{11} , the velocity tends to that of Rayleigh waves in the layer ($c'_R = 0.92\beta'$), whereas for the other modes the limit is the S wave velocity in the layer (β').

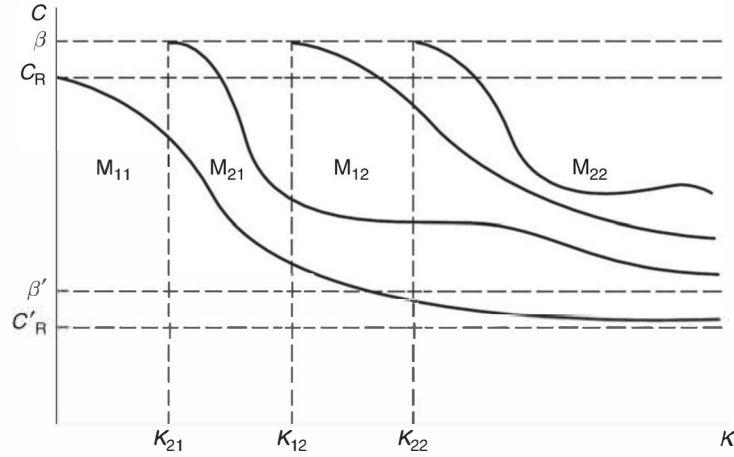


Fig. 12.14

Dispersion curves of Rayleigh waves in an elastic layer over a half-space for the fundamental mode (M_{11}) and the three first higher modes (symmetric M_{12} , and antisymmetric M_{21} and M_{22}).

12.5 Surface waves in layered media

The matrix formulation for the propagation of body waves in layered media, as we have seen in [Chapter 7](#), can be applied to the propagation of surface waves (Haskell, 1953). This constitutes a very efficient method for the numerical determination of the dispersion curves of Rayleigh and Love waves in multilayered media with constant parameters. For Love waves we use the formulation for SH motion ([Sections 7.4.2 and 7.6.1](#)). The dispersion equation for Love waves in a medium of N layers is found by imposing the condition of transmission of SH waves in the layers and a decrease with depth in amplitude in the half-space (s_N is imaginary and $B_N = 0$). The dispersion equation ($f(\omega, c) = 0$) corresponds to the zeroth value of the amplitude in the half-space. For Rayleigh waves we use the formulation for the P-SV motion ([Section 7.6.2](#)). As in the case of Love waves, imposing the conditions that waves propagate only in the x direction in the half-space and their amplitudes decrease with depth (r_N and s_N have imaginary values and $B_N = D_N = 0$), we obtain the dispersion equation ($f(c, \omega) = 0$).

12.5.1 Love waves in a layer over a half-space

As an example we apply the matrix formulation to find the dispersion equation for Love waves in a layer over a half-space. This is the problem we solved in [Section 12.3](#). We begin with formulations for the SH motion in a layer over a half-space from [Section 7.5](#). The relation between the amplitudes of motion at the layer and at the half-space is given by equation (7.64). From this equation we can derive the dispersion equation of Love waves, imposing the condition that there are no SH waves propagating in the z direction in

the half-space; that is, s_2 is imaginary ($c < \beta_2$). Since amplitudes must decrease with z positive, it follows also that $B_2 = 0$. With these conditions, the equation resulting from the zeroth component of the vector $b_1(0)$ in (7.64) is

$$A_2(a_1 \sin d_1 e^{id_2} + ia_2 \cos d_1 e^{id_2}) = 0. \quad (12.94)$$

The existence of Love waves which propagate in the x direction and attenuate in the z direction in the half-space requires that A_2 is not zero. The expression inside the brackets must then be zero and we obtain

$$\frac{\sin d_1}{\cos d_1} = -\frac{ia_2}{a_1}. \quad (12.95)$$

If we substitute the values for a_1 , d_1 , a_2 , and d_2 (below 7.61), and put $s_2 = i\bar{s}_2$, we obtain

$$\tan(kHs_1) = \frac{\mu_2 \bar{s}_2}{\mu_1 s_1}.$$

This is the same dispersion equation (12.60) we have already derived for Love waves. As we saw in Section 12.3, there is no solution for s_1 imaginary. Since s_1 is real, SH waves propagate inside the layer in both z directions (positive and negative). These waves are totally reflected from the contact surface so that there are only inhomogeneous waves (s_2 imaginary) in the half-space.

12.6 Surface waves in a spherical medium

In a flat medium, surface waves propagate along the surface and have cylindrical symmetry. Using the solution of the wave equation in cylindrical coordinates (5.129), for symmetry with respect to ϕ , we can write the displacements as

$$u(R, t) = \left(\frac{2}{\pi k R}\right)^{1/2} A \exp\left[i\left(kR - \omega t - \frac{\pi}{4}\right)\right]. \quad (12.96)$$

The amplitude depends on the square root of the distance R along the surface. Thus the energy is constant along a circle of radius $2\pi R$.

In a spherical medium of radius a , the displacements of surface waves, without dependence on r , have the form of the solutions of the wave equation in spherical coordinates (5.111). For symmetry with respect to ϕ , they can be written as

$$u(\theta, t) = A_n \left(\frac{2}{\pi k}\right)^{1/2} \frac{1}{a} P_n(\cos \theta) \exp\left[i\left(\omega t - (n+1)\frac{\pi}{2}\right)\right]. \quad (12.97)$$

For wave lengths that are small relative to the radius, they correspond to high values of n and we can use the asymptotic expansion of $P_n(\cos \theta)$, which, for values of θ away from the poles ($\theta = 0$ and π), is given by

$$P_n(\cos \theta) \simeq \left(\frac{2}{2n \sin \theta} \right)^{1/2} \cos \left[\left(n + \frac{1}{2} \right) \theta - \frac{\pi}{4} \right]. \quad (12.98)$$

Then, the displacements of surface waves may be written in a simplified way in terms of the angular distance Δ (Ben Menahem and Singh, 1981):

$$u_s \simeq A \left(\frac{\pi}{\sin \Delta} \right)^{1/2} \exp \left[i \left(ka\Delta - \omega t - \frac{3\pi}{4} \right) \right]. \quad (12.99)$$

Similarly to the flat case, the amplitudes depend on the inverse of the square root of the distance ($\sin \Delta$) and the energy is constant on a circle of radius $2\pi \sin \Delta$. Surface waves in a sphere can be approximated by those of a flat medium if their wave length is small compared with the radius. The wave number corresponding to a spherical medium k_s can be obtained from that of the flat medium k_f , by introducing the correction (Ben Menahem and Singh, 1981)

$$k_s = \left(k_f^2 + \frac{9}{4a^2} \right)^{1/2}. \quad (12.100)$$

From this expression we can find the relations between the phase velocities of the spherical and flat cases, c_s (spherical) and c_f (flat) ($c = \omega/k$),

$$c_s = c_f \left(1 + \frac{9c_f^2}{4a^2\omega^2} \right)^{-1/2}, \quad (12.101)$$

and group velocities U_s and U_f ($U = d\omega/dk$),

$$U_s = U_f \left(1 + \frac{9}{4a^2k_f^2} \right)^{1/2}. \quad (12.102)$$

An important phenomenon in the propagation of surface waves in a sphere is the polar phase shift of $\lambda/4$ when waves cross the poles. This is due to the fact that harmonic waves are not exactly sinusoidal or cosinusoidal in the vicinity of the poles. This does not affect the determination of group velocities but affects phase velocities. Thus, for determination of phase velocities along paths that include the epicenter or its antipole, a phase shift of $\pi/2$ must be added for each pole or antipole crossing. This effect must be taken into account when determining phase velocities using surface waves that have circled the Earth (Brune *et al.*, 1961).

12.7 Stoneley waves

We have seen how the presence of boundaries between layers of different characteristics affect the generation of surface waves. We consider now the problem of whether there exist

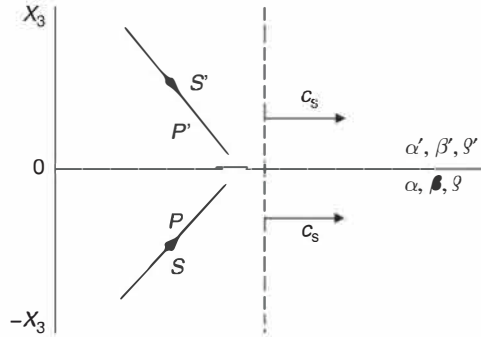


Fig. 12.15

Waves incident on the surface of contact between two elastic half-spaces. Stoneley waves are shown.

waves similar to surface waves related to the contact surface of two media without the presence of a free surface. Let us consider the existence of such waves in two elastic half-spaces of different characteristics, in contact (Fig. 12.15). These waves are generated by constructive interference of an incident wave on the contact surface and their amplitudes must decrease exponentially with the distance from the contact surface in both half-spaces. Their existence was shown by Robert Stoneley (1924), and they are known by his name. These waves cannot be observed on seismographs installed on the free surface. The potentials ϕ and ψ and displacements u_2 for waves propagating in the x_1 positive direction and with amplitudes decreasing with distance from the contact surface ($x_3 = 0$) in each medium are given by

$$\phi = A \exp[-ikrx_3 + ik(x_1 - ct)], \quad (12.103)$$

$$\psi = B \exp[-iksx_3 + ik(x_1 - ct)], \quad (12.104)$$

$$u_2 = C \exp[-iksx_3 + ik(x_1 - ct)], \quad (12.105)$$

$$\phi' = A' \exp[ikr'x_3 + ik(x_1 - ct)], \quad (12.106)$$

$$\psi' = B' \exp[iks'x_3 + ik(x_1 - ct)], \quad (12.107)$$

$$u'_2 = C' \exp[iks'x_3 + ik(x_1 - ct)]. \quad (12.108)$$

The boundary conditions are the continuity of stress and displacement components across the contact surface:

$$\begin{aligned} \tau_{31} &= \tau'_{31}; & \tau_{32} &= \tau'_{32}; & \tau_{33} &= \tau'_{33} \\ u_1 &= u'_1; & u_2 &= u'_2; & u_3 &= u'_3 \end{aligned}.$$

We proceed as in the [previous section](#), writing stresses as functions of displacements and these in terms of potentials. The conditions for τ_{32} and u_2 give the result $C = C' = 0$. As in the case of Rayleigh waves in a half-space, Stoneley waves have no transverse component of displacement. From the other four conditions we obtain

$$A + Bs = A' - B's', \quad (12.109)$$

$$-rA + B = r'A' + B', \quad (12.110)$$

$$\beta^2 \rho [2rA - (1 - s^2)B] = \beta'^2 \rho [2r'A - (1 - s'^2)B'], \quad (12.111)$$

$$\rho [\alpha^2 (r^2 + 1) - 2\beta^2] A - 2\beta^2 \rho s B = \rho' [\alpha'^2 (r'^2 + 1) - 2\beta'^2] A' - 2\beta'^2 s' \rho' B'. \quad (12.112)$$

The condition for the existence of a solution is that the determinant of the system is null. For the particular case in which $\sigma = \frac{1}{4}(\lambda = \mu)$ in both media, the determinant is given by

$$\begin{vmatrix} 1 & s & -1 & s' \\ -r & 1 & -r' & -1 \\ 2r\mu & -(1 - s^2)\mu & -2r'\mu' & (1 - s'^2)\mu \\ \mu(3r^2 - 1) & -2\mu s & -\mu'(3r'^2 - 1) & 2\mu's' \end{vmatrix} = 0.$$

By expanding the determinant we obtain a fourth-order equation for c . The four roots are real only for certain values of the ratios ρ/ρ' and β/β' . If β' is less than β , the velocity of Stoneley waves c_s is in the range $\beta' < c_s < c_R$, where c_R is the Rayleigh wave velocity for the half-space of velocity β . Stoneley waves have characteristics similar to Rayleigh waves, with prograde or retrograde elliptical particle motion depending on the relative characteristics of the two media and generally a vertical major axis. The frequency does not appear in equations (12.109)–(12.112); therefore, waves are not dispersed and their velocity is constant, just like for Rayleigh waves in a half-space.

12.8 Summary

The presence on the Earth of a free surface and of layers with different elastic properties leads to the problem of the generation of surface waves. First, on an elastic half-space of constant velocity with a free surface, Rayleigh waves are produced with constant velocity, a fraction of the shear velocity in the medium, and no transverse component of horizontal motion. The presence of layers with different elastic properties in addition to the free surface produces surface waves with both radial and transverse horizontal components with different velocities (Love and Rayleigh waves); in both cases velocities depend on frequency, presenting the phenomenon of wave dispersion and the presence of modes of propagation. These waves are produced by the constructive interference of supercritical incident and reflected waves on the layer boundaries.

The problem has been studied in detail for a single layer over a half-space. Love waves with only a horizontal transverse component are produced, which present wave dispersion, fundamental and higher modes, and velocities between the values of the

shear waves of the medium and the layer. In the half-space, displacement is maximum at the surface and decreases with depth. Inside the layer there are nodes of the displacement for the higher modes, not for the fundamental. Rayleigh waves have radial horizontal and vertical components and also present dispersion. Higher modes are now divided into symmetric and antisymmetric. The general problem of surface wave propagation on a multilayer medium can be solved by applying the matrix formulation we saw in Chapter 7. An example of this application has been given for the problem of Love waves on a single layer. Finally we have presented briefly the problem of surface waves on a spherical medium and the generation of Stoneley waves on the interface between two half-spaces.

12.9 Problems

- P12.1 A Rayleigh wave in a semi-infinite medium has a 20 s period. If the P wave velocity is 6 km/s and Poisson's ratio 0.25, calculate the depth at which $u_1 = 0$, and at which depth the particle movement becomes prograde.
- P12.2 There is a liquid layer of density ρ and speed of propagation α on top of a rigid medium (half-space). Derive the dispersion equation of waves in the layer by boundary conditions and by constructive interference in terms of ω . Plot the dispersion curve for the different modes.
- P12.3 Consider a liquid layer of thickness H , density ρ' , and velocity α' over an elastic half-space for which $\lambda = \mu$. Write the potentials ϕ and ψ for the layer and half-space, the boundary conditions, and the determinant that gives the equation for the velocity of surface waves.
- P12.4 Consider an elastic layer of coefficients λ and μ , thickness H , and density ρ on a rigid semi-infinite medium. Derive the dispersion equation $c(\omega)$ for P-SV and SH type channeled waves for the fundamental mode (FM) and the first higher mode (1HM). Plot the dispersion curve for the SH motion.
- P12.5 Consider an elastic layer of thickness H , coefficients λ and μ , and density ρ over a rigid half-space. Deduce the dispersion equation $c(\omega)$ for guided motion of waves of SH and P-SV type for the fundamental mode and the first higher mode. Draw the dispersion curve for the SH wave motion.
- P12.6 Consider a layer of thickness H , density ρ , and rigidity μ' over a half-space of density ρ and rigidity μ , with the relations $\mu = 4\mu'$, $\rho = \rho'$, $\alpha = c/\sqrt{2}$, and $b = H/\lambda$.
- Write the dispersion equation of Love waves in terms of a and b .
 - Find the values of b for $a = 1$, $\frac{3}{4}$, and $\frac{1}{2}$; and plot a versus b .
- P12.7 Given a layer of thickness H and shear modulus $\mu = 0$ on top of a half-space or semi-infinite medium in which $\lambda = 0$, study (without expanding the determinant) whether there exist surface waves that propagate in the direction x_1 . Are they dispersive waves?

We have seen that surface waves in layered media are dispersed; that is, their velocity is a function of the frequency (or period). Thus, for an impulsive time function at the source, surface waves at some distance are formed by trains of waves, different frequencies arriving at different times. This is an important phenomenon that requires further consideration. Arrival times, amplitudes, and phases for each frequency depend, then, on the dispersion equation. In [Section 5.4](#) we saw that, if the phase velocity is a function of the frequency, then the velocity of energy transport is not the same, but equal to the group velocity, or the velocity of propagation of wave groups. We will consider now, with more detail, the general problem of wave dispersion and the relation between phase and group velocities and apply it to surface waves generated by earthquakes in the Earth. The study of the dispersion curves of phase and group velocities of surface waves is used to determine the structure of the Earth's crust and mantle.

13.1 Phase and group velocities

The displacement of a sinusoidal wave of angular frequency ω and wave number k that propagates in the x direction is given by

$$u(x, t) = A \sin[(kx - \omega t) + \phi], \quad (13.1)$$

where the phase velocity, or the velocity of propagation of each value of the phase, is

$$c = \omega/k. \quad (13.2)$$

For monochromatic waves in a homogeneous medium, c is constant and for each value of ω there is a single value of k . In this case, the velocity of energy transport is equal to the phase velocity ([Section 5.4](#)); this is the case for body waves. If the phase velocity is a function of the frequency $c(\omega)$, then we can also write $k(\omega)$ and $\omega(k)$, and we can use as the independent variable either k or ω . In the first case we are looking at the wave phenomenon from the point of view of its dependence on space ($k = 2\pi/\lambda$) and in the second, in terms of its dependence on time ($\omega = 2\pi/T$). As we saw in [Section 5.4](#), the group velocity is given by

$$U = d\omega/dk. \quad (13.3)$$

In (13.2), putting $\omega = ck$, if we take derivatives with respect to k we obtain

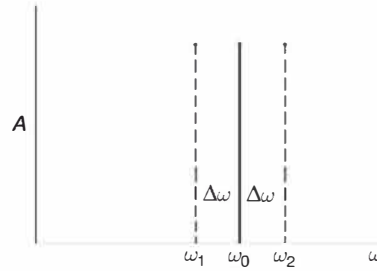


Fig. 13.1 Two frequencies at an interval $\Delta\omega$ from a central ω_0 .

$$U = c + k \frac{dc}{dk}. \quad (13.4)$$

For the dependence on ω , taking the derivative with respect to ω , we obtain

$$U = \frac{c}{1 - \frac{\omega}{c} \frac{dc}{d\omega}}. \quad (13.5)$$

The simplest case of a wave with more than one frequency is that formed by the sum of two waves with the same amplitude and two similar frequencies, namely, $\omega_1 = \omega_0 - \Delta\omega$ and $\omega_2 = \omega_0 + \Delta\omega$ (Fig. 13.1). The displacement for null initial phase, $\phi = 0$, is

$$u = A \sin(k_1 x - \omega_1 t) + A \sin(k_2 x - \omega_2 t).$$

This equation can be written as

$$u = 2A \sin\left(\frac{k_1 + k_2}{2}x - \frac{\omega_1 + \omega_2}{2}t\right) \cos\left(\frac{k_1 - k_2}{2}x - \frac{\omega_1 - \omega_2}{2}t\right), \quad (13.6)$$

and, on substituting for k_1 , k_2 , ω_1 , and ω_2 in terms of ω_0 and $\Delta\omega$,

$$u = 2A \sin(k_0 x - \omega_0 t) \cos(\Delta k x - \Delta\omega t). \quad (13.7)$$

The displacement has the form of a sine wave of frequency ω_0 modulated by a cosine wave of frequency $\Delta\omega$. It is, then, formed by groups or packets of waves of frequency ω_0 with duration $\pi/\Delta\omega$ (Fig. 13.2). According to (13.2), the phase velocity of sine waves is $c = \omega_0/k_0$, whereas the maxima of the cosine function propagate with velocity $U = \Delta\omega/\Delta k$. Since groups of waves correspond to maxima of the cosine function, this is the group velocity. Energy is associated with amplitude maxima and is propagated by the group velocity U . Then, even for a wave formed by two frequencies we can distinguish between the phase and group velocities.

A wave with a discrete content of N frequencies of values ω_k can be represented by the sum of the displacements of each of these frequencies:

$$u(x, t) = \sum_{k=1}^N A_k \cos\left[\omega_k \left(\frac{x}{c_k} - t\right) + \phi_k\right], \quad (13.8)$$

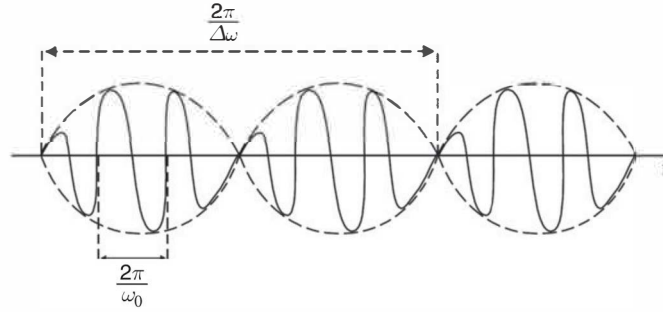


Fig. 13.2 The sum of two waves of frequencies at an interval $\Delta\omega$ from a central ω_0 .

where c_k is the phase velocity corresponding to each frequency ω_k , A_k is the amplitude, and ϕ_k is the initial phase. If the content of frequencies is continuous between zero and infinity, then the displacement is expressed by the integral:

$$u(x, t) = \int_0^\infty A(\omega) \cos \left[\omega \left(\frac{x}{c(\omega)} - t \right) + \phi(\omega) \right] d\omega. \quad (13.9)$$

$A(\omega)$ is the amplitude corresponding to each frequency and $\phi(\omega)$ is the initial phase. The wave displacement $u(x, t)$ is the sum of the contributions from all frequencies. The displacement observed at a particular distance x is a function of time $u(t)$ (a seismogram) and by means of Fourier transformation can be expressed as a function of frequency $u(\omega)$ (a complex spectrum) (Section 5.2.2; Appendix 4).

13.2 Groups of waves

Let us consider now a wave formed by a continuous distribution of frequencies limited to a narrow band centered at a frequency ω_0 and a bandwidth $2\Delta\omega$ (Fig. 13.3a). Using as a variable the wave number k , the displacements from equation (13.9), with zero initial phase for all frequencies and constant amplitude A_0 , are given by

$$u(x, t) = A_0 \int_{k_0 - \Delta k}^{k_0 + \Delta k} \cos(kx - \omega t) dk. \quad (13.10)$$

On performing a *Taylor series expansion* (first proposed by Brook Taylor in 1715) of the phase around k_0 and taking only the term with the first derivative, we have

$$kx - \omega t = k_0 t + (k - k_0) \frac{d}{dk} (kx - \omega t)_{k_0} + \dots \quad (13.11)$$

On substituting into (13.10), we have

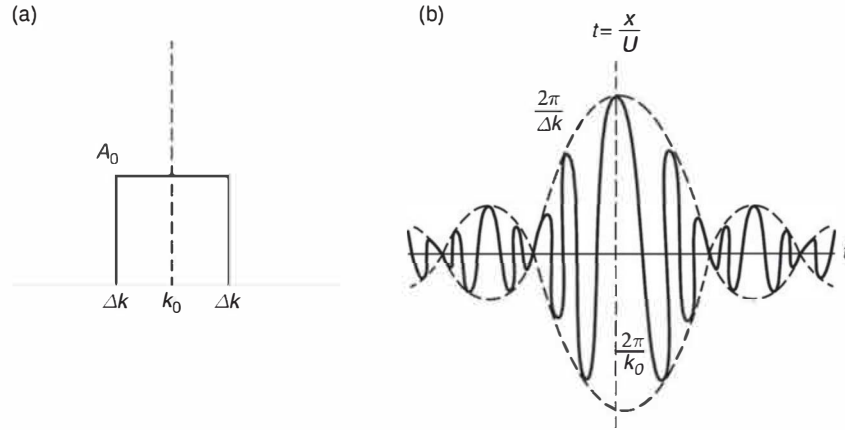


Fig. 13.3

A wave formed by a narrow band of frequencies centered at k_0 and of width $2\Delta k$. (a) The band of frequencies. (b) The wave as a function of time, the product of $\cos(\omega_0 t)$ and $\sin X/X$.

$$u = A_0 \int_{k_0 - \Delta k}^{k_0 + \Delta k} \cos \left[kx_0 - \omega t_0 - \left(x - \frac{d\omega}{dk} t \right) k_0 + \left(x - \frac{d\omega}{dk} t \right) k \right] dk. \quad (13.12)$$

The derivatives $d\omega/dk$ are evaluated at $k = k_0$. This integral can be written as

$$u = A_0 \int_{k_0 - \Delta k}^{k_0 + \Delta k} \frac{1}{b} \cos(a + bk) dk, \quad (13.13)$$

$$a = k_0 \frac{d\omega}{dk} t - \omega_0 t = k_0 U t - \omega_0,$$

$$b = x - \frac{d\omega}{dk} t = x - U t,$$

where $U(k) = \frac{d\omega}{dk}$ is the group velocity. On solving the integral,

$$u = \frac{A_0}{b} \{ \sin[a + b(k_0 + \Delta k)] - \sin[a + b(k_0 - \Delta k)] \}. \quad (13.14)$$

Since $\sin(x + y) - \sin(x - y) = 2 \sin x \cos y$, we obtain

$$u = \frac{2A_0}{b} \sin(\Delta k b) \cos(a + b k_0). \quad (13.15)$$

On replacing the values of a and b and multiplying and dividing by Δk , we have

$$u(x, t) = A_0 \Delta k \frac{\sin X}{X} \cos(k_0 x - \omega_0 t), \quad (13.16)$$

$$X = \left(x - \frac{d\omega}{dk} t \right) \Delta k = (x - U t) \Delta k. \quad (13.17)$$

The displacement $u(x, t)$ is a cosine wave of wave number k_0 , modulated by the function $\sin X/X$. This function has a maximum value equal to unity for $X = 0$, and zeros for $X = \pm\pi, \pm2\pi$, etc. The main pulse is centered at $X = 0$, and the amplitudes of the other pulses decrease with $1/X$. As a function of time for a fixed distance, the displacement is formed by a group or packet, of width $2\pi/\Delta\omega$, of waves of frequency ω_0 . The maximum of the wave group arrives at $t = x/U$ (Fig. 13.3b). The group velocity U is, then, the velocity of propagation of the wave packet formed by the envelope of the function $\sin X/X$. The phase velocity of elementary waves contained in the packet with frequency ω_0 is $c = \omega_0/k_0$.

13.3 The principle of a stationary phase

The principle of a *stationary phase* is used in order to evaluate displacements of dispersed waves with a continuous distribution of frequencies with a broad spectrum (13.9). As a function of the wave number k and assuming all initial phases to be zero, equation (13.9) can be written as

$$u(x, t) = \int_{-\infty}^{\infty} A(k) \cos[\Phi(k)] dk, \quad (13.18)$$

where

$$\Phi(k) = kx - \omega t. \quad (13.19)$$

If the amplitude $A(k)$ is a slowly varying function of k compared with the variation of the phase $\Phi(k)$, the integral (13.18) has significant values only for frequencies for which $\Phi(k)$ is stationary; that is, when $d\Phi/dk = 0$ (Fig. 13.4). For rapid variations of $\Phi(k)$, the cosine function changes sign and its integral becomes null, except for those values of $k = k_0$, for which $\Phi(k)$ does not vary ($d\Phi/dk = 0$). We call k_0 and ω_0 the wave number and

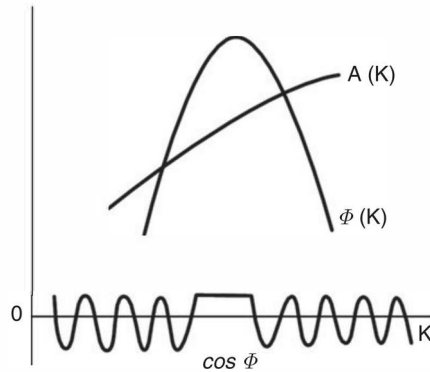


Fig. 13.4

The shapes of functions $A(k)$, $\Phi(k)$, and $\cos \Phi$, showing the contribution of the stationary phase.

frequency for which the phase is stationary. The integral (13.18) becomes, in a similar way to that in (13.10),

$$u(x, t) = A(k_0) \int_{k_0 - \Delta k}^{k_0 + \Delta k} \cos[\Phi(k)] dk. \quad (13.20)$$

This integral has values distinct from zero only for values of k near k_0 .

If we take the derivative with respect to k of Φ and put it equal to zero,

$$\frac{d\Phi}{dk} = \frac{d}{dk}(kx - \omega t) = 0; \quad (13.21)$$

then, since this is zero for $k = k_0$, and $d\omega/dk = U$, we obtain

$$U(k_0) = \frac{x}{t}, \quad (13.22)$$

where $U(k_0)$ is the group velocity corresponding to the wave number k_0 . Then, at a distance x and time t , energy is contained in waves with wave number k_0 or frequency ω_0 that correspond to stationary values of the phase. According to the principle of a stationary phase, for that distance and time, the other frequencies do not contribute to the amplitude. For each value of x and t , k_0 has a different value. This result is similar to that of the [previous section](#), in which k_0 was the central value of the wave-number band. Here we have all wave numbers but, for each pair of values of x and t , k_0 is the only value that makes the phase stationary.

An equivalent expression to (13.21) is obtained by taking derivatives of the phase Φ with respect to the frequency ω :

$$\frac{d\Phi}{d\omega} = \frac{d}{d\omega}(kx - \omega t) = \frac{x}{U} - t. \quad (13.23)$$

For the stationary phase, $d\Phi/d\omega = 0$, and for the corresponding frequency ω_0 , we obtain the same result as that in (13.22). The group velocity is, then, the velocity corresponding to the frequency that makes the phase stationary for a given distance and time. If we write Φ in terms of the phase velocity, the condition of the stationary phase results in

$$\frac{d\Phi}{d\omega} = \frac{d}{d\omega} \left[\omega \left(\frac{x}{c} - 1 \right) \right] = 0, \quad (13.24)$$

and, since for the stationary phase $x/t = U$, from (13.24) we obtain the relation between the group and phase velocities (13.5):

$$U = \left(\frac{1}{c} - \frac{\omega}{c^2} \frac{dc}{d\omega} \right)^{-1}. \quad (13.25)$$

In order to determine the amplitudes corresponding to the stationary phase, we start with the equation

$$u(x, t) = \int_{-\infty}^{\infty} A(k) e^{i(kx - \omega t)} dk, \quad (13.26)$$

where, for simplicity, we have assumed the initial phases to be zero. According to the principle of the stationary phase, this integral is null outside the values near the frequency which makes the phase stationary (13.20). On performing a Taylor expansion of the phase about the wave number k_0 , including the second derivatives, we have

$$kx - \omega t = k_0x - \omega_0t + (k - k_0) \frac{d}{dk} [kx - \omega t]_{k=k_0} + \frac{1}{2} (k - k_0)^2 \frac{d^2}{dk^2} [kx - \omega t]_{k=k_0}. \quad (13.27)$$

As we have seen, for the stationary phase, the term of the first derivative is null. On substituting (13.27) into (13.26), we obtain

$$u(x, t) = A(k_0) e^{i(k_0x - \omega_0t)} \int_{-\infty}^{\infty} \exp \left(i \frac{1}{2} (k - k_0)^2 \frac{dU}{dk} t \right) dk, \quad (13.28)$$

where we have used that

$$\frac{d}{dk} (kx - \omega t) = x - Ut. \quad (13.29)$$

On making the change of variable,

$$\sigma^2 = \frac{1}{2} (k - k_0)^2 \frac{dU}{dk} t, \quad (13.30)$$

equation (13.28) takes the form,

$$u(x, t) = A(k_0) e^{i(k_0x - \omega_0t)} \left[\frac{t}{2} \frac{dU}{dk} \right]_{k_0}^{-1/2} \int_{-\infty}^{\infty} e^{-i\sigma^2} d\sigma. \quad (13.31)$$

The definite integral has the value $(i\pi)^{1/2}$. On substituting into (13.31), since $\sqrt{i} = \pm \exp(i\pi/4)$, we obtain for the real part of (13.31),

$$u(x, t) = A(k_0) \left(\frac{2\pi}{\frac{x}{U} \frac{dU}{dk}} \right)^{1/2} \cos \left(k_0x - \omega_0t \pm \frac{\pi}{4} \right). \quad (13.32)$$

Then, for a given time and distance, energy is contained in a cosine wave of wave number k_0 and frequency ω_0 , values corresponding to the stationary phase. Since dU/dk is in the denominator, the largest amplitudes correspond to $dU/dk = 0$, or the *Airy phase*, first proposed by Airy for gravity waves in the ocean. In this case, in the series expansion (13.27), the second term is now zero and we must take the third term. The amplitudes of displacement of the Airy phase are given by

$$u(x, t) = A(k_0) \left(\frac{2\pi}{\frac{x}{U} \frac{d^2U}{dk^2}} \right)^{1/3} \cos \left(k_0x - \omega_0t \pm \frac{\pi}{4} \right). \quad (13.33)$$

An example of the distribution of amplitudes as a function of time in a dispersed train of surface waves is shown in Fig. 13.5. The phase velocity curve $c(k)$ is that of

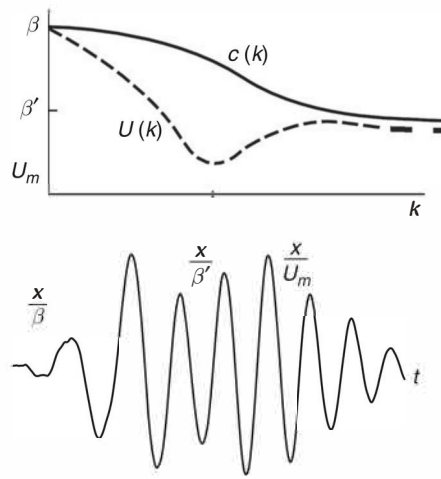


Fig. 13.5 Phase- and group-velocity dispersion curves and the corresponding train of dispersed waves at a distance x as functions of time.

Love waves for a layer over a half-space (Section 12.3). The group velocity curve $U(k)$ is derived from the phase velocity according to (13.4) (Fig. 13.5 top). With this curve and equation (13.32), amplitudes observed at a distance x as a function of time are shown in Fig. 13.5 bottom. The waves observed are those corresponding to the frequencies that make the phase stationary. In consequence, they travel with the group velocity. Waves corresponding to the lowest frequencies arrive first, starting at time $t = x/\beta$. This fact makes it easy to identify the beginning of the surface waves train on the seismograms (Fig. 11.21). At time $t = x/\beta'$, waves with the highest frequencies that are superimposed on those of lower frequencies arrive. The last arrival, at $t = x/U_m$, is that of waves that travel with the minimum group velocity. At this time, waves with similar frequencies near k_m arrive together and are summed up to form the Airy phase. According to (13.32), on a seismogram recorded at a given distance, for each value of time, we observe only the wave corresponding to the frequency that makes the phase stationary.

13.4 Characteristics of dispersed waves

At a given distance, dispersed waves have the form of trains of waves with different frequencies arriving at different times, according to the frequency dependence of the velocity (the dispersion curve). Since the frequencies observed are those corresponding to stationary values of the phase, their time of arrival depends on the group velocity rather than on the phase velocity. This is an important aspect of dispersion, which is not always well understood, that can be further explained by considering the derivatives of the phase with respect to distance and time.

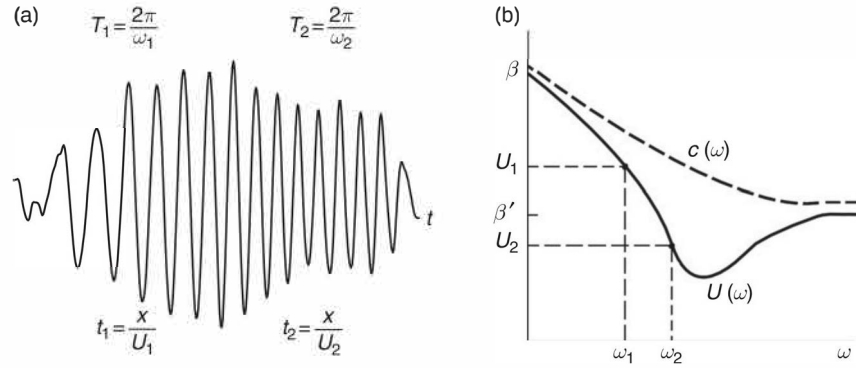


Fig. 13.6

(a) Arrivals of waves of instantaneous frequencies ω_1 and ω_2 in a dispersed wave train. (b) Phase- and group-velocity dispersion curves.

Taking derivatives of the phase Φ with respect to distance for a fixed time, we obtain

$$\frac{\partial}{\partial x}(kx - \omega t) = k + \frac{\partial k}{\partial x}(x - Ut). \quad (13.34)$$

For a wave number k_0 that corresponds to a stationary value of the phase, $U(k_0) = x/t$ and, consequently,

$$\frac{\partial \Phi}{\partial x} = k_0. \quad (13.35)$$

This means that, for a fixed time, at a distance x , energy arrives in a wave with wave number k_0 , called the *local wave number*, which is different for every distance. Its wave length is $\lambda_0 = 2\pi/k_0$.

If we take the time derivative of the phase Φ , for a fixed distance, we obtain

$$\frac{\partial}{\partial t}(kx - \omega t) = -\omega + \frac{\partial k}{\partial t}(x - Ut). \quad (13.36)$$

Just like the preceding case with the frequency ω_0 corresponding to the stationary phase, we have

$$\frac{\partial \Phi}{\partial t} = -\omega_0. \quad (13.37)$$

For a fixed distance, at each time, energy arrives in a wave of frequency ω_0 , called the *instantaneous frequency*, which is different for every time. Its period is $T_0 = 2\pi/\omega_0$, and, for a fixed distance, in the train of dispersed waves, at each time, we observe only this period (Fig. 13.6). In conclusion, waves observed at a given distance and time correspond to the instantaneous frequency ω_0 or to the *local wave number* k_0 , and propagate with the group velocity.

Let us consider now an increment in phase $\delta\Phi$, related to changes in distance and time and also in frequency and wave number:

$$\delta(kx - \omega t) = k \delta x + x \delta k - t \delta \omega - \omega \delta t. \quad (13.38)$$

For frequencies and wave numbers corresponding to stationary values of the phase $x \delta k - t \delta \omega = 0$, we obtain

$$\delta \Phi = k_0 \delta x - \omega_0 \delta t. \quad (13.39)$$

If we observe the same phase in a wave of the same frequency ω_0 , at two points separated by a distance δx , with a time interval δt , then the phase increment is null ($\delta \Phi = 0$) and we can find that

$$\frac{\delta x}{\delta t} = \frac{\omega_0}{k_0} = c, \quad (13.40)$$

where c is the phase velocity of waves with frequency and wave number ω_0 and k_0 . This expression should be compared with equation (13.22) for the group velocity. Using equations (13.22) and (13.40) we can determine both group and phase velocities from observations of dispersed wave trains. Consider a wave observed at distance x and time t , with instantaneous frequency ω_0 ; we obtain the group velocity $U(\omega_0) = x/t$. Consider now a wave observed at two nearby points separated by a distance δx . If the same phase for the same instantaneous frequency ω_0 arrives with a time difference δt , we can calculate the phase velocity $c(\omega_0) = \delta x / \delta t$. These ideas are at the base of methods used for determination of phase and group velocities from observations of dispersed wave trains.

13.5 Determination of group and phase velocities. Instantaneous frequencies

The determination of group and phase velocities from observations of dispersed surface waves is based on ideas discussed previously. A seismogram is the recording of waves as a function of time for a given distance. The amplitudes of dispersed surface waves, according to (13.32), are given by

$$u(x, t) = A(\omega_0) \left(\frac{2\pi}{\frac{x}{U} \frac{dU}{dk}} \right)^{1/2} \cos \left(k_0 x - \omega_0 t + \phi + \phi_I \pm \frac{\pi}{4} \right), \quad (13.41)$$

where we have introduced the initial phase ϕ and the phase shift ϕ_I produced by the instrument. For a given distance x (the epicentral distance), a wave of instantaneous frequency ω_0 arrives at time t . As has already been explained, frequencies observed directly in the seismograms are the frequencies that correspond to stationary values of the phase. Methods that use directly recorded waves are, then, limited to these frequencies.

13.5.1 The group velocity

The group velocity can be determined from the record of surface waves at a single station. If we correct for the instrumental phase shift, ϕ_I , the phase is given by

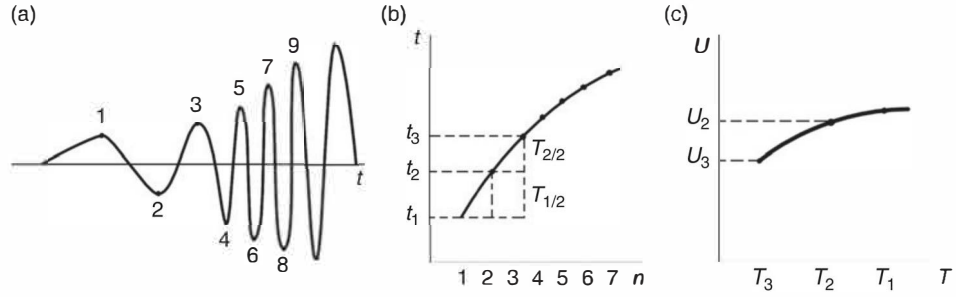


Fig. 13.7

Determination of the group velocity from the instantaneous frequencies at one station. (a) Identification of peaks and troughs. (b) Traveling times of peaks and troughs and determination of their period. (c) The group velocities corresponding to each period.

$$\Phi = kx - \omega t + \phi \pm \frac{\pi}{4}. \quad (13.42)$$

On taking the derivative with respect to ω and, since $d\omega/dk = U$, solving for U , we obtain

$$U = \frac{x}{\frac{d\Phi}{d\omega} + \frac{d\phi}{d\omega} + t}, \quad (13.43)$$

for a stationary phase $d\Phi/d\omega = 0$ and $\omega = \omega_0$. Assuming that the initial phase does not depend on frequency, we obtain for each instantaneous frequency ω_0 ,

$$U(\omega_0) = \frac{x}{t(\omega_0)}. \quad (13.44)$$

The method for determination of the group velocity consists of measuring times of arrival of peaks and troughs of waves in a dispersed train (these are phases $\Phi = 0$ and π) (Fig. 13.7a). These values are represented in a plot with respect to the order number (Fig. 13.7b). On doubling the intervals between pairs of values (or from peak to peak) we obtain the periods corresponding to instantaneous frequencies and, from the ordinates, we obtain their arrival times. On dividing the epicentral distance by each arrival time (13.44), we obtain the group velocity $U(T_0)$ for each period (Fig. 13.7c). This velocity corresponds to a mean value of the velocities of the structure along the trajectory from the epicenter to the station.

13.5.2 The phase velocity

On a seismogram of dispersed surface waves, peaks correspond to values of the phase equal to multiples of 2π . The waves recorded are those corresponding to instantaneous frequencies (13.41), so that we can write

$$k_0 x - \omega_0 t + \phi + \phi_f \pm \frac{\pi}{4} = 2N\pi, \quad (13.45)$$

where $N = 0, 1, 2, 3, \dots$

After correcting for the instrumental phase, ϕ_I , on dividing (13.45) by k_0 , we obtain for the phase velocity ($c = \omega/k$),

$$c(T_0) = \frac{x}{t - (\phi + N \pm \frac{1}{8})T_0}. \quad (13.46)$$

If we want to determine c , we need to know the value of the initial phase ϕ , corresponding to the waves arriving at the station for each instantaneous frequency. This is possible only if we know the focal mechanism of the earthquake with the distribution of initial phases around the focus. Only in this case can we use equation (13.46) to determine the phase velocity from the focus to the station. The method consists of selecting the peaks (or peaks and troughs) in the seismogram (phases $\Phi = 2N\pi$) and measuring their traveling times t_i and periods T_i . For the same periods we determine from the focal mechanism the initial phase ϕ , corresponding to the waves arriving at the station and substitute them into equation (13.46). Values of N are found by trial and error, by successive substitutions of 0, 1, 2, etc., selecting those which give a reasonable value of the velocity. The value of N corresponds to the number of complete cycles that must be added to the phase for each period so that they correspond to the observed phase.

The necessity of knowing the focal mechanism if one is to calculate the phase velocity from a single station is the reason why this method is rarely used. This difficulty is avoided by using two stations with equal instruments. In this method two stations that are lined up with the epicenter are selected (that is, both stations and the epicenter lie on the same great circle). For the same instantaneous frequency ω_0 , the phases corresponding to peaks at each station are given by: station 1 at distance x_1

$$k_0 x_1 - \omega_0 t_1 + \phi + \phi_I \pm \frac{\pi}{4} = 2L\pi; \quad (13.47)$$

station 2 at distance x_2

$$k_0 x_2 - \omega_0 t_2 + \phi + \phi_I \pm \frac{\pi}{4} = 2M\pi. \quad (13.48)$$

We subtract (13.47) from (13.48) and solve for $c = \omega/k$, putting $\Delta x = x_2 - x_1$, $\Delta t = t_2 - t_1$, and $N = M - L$, and obtain

$$c(T_0) = \frac{\Delta x}{\Delta t - NT_0}. \quad (13.49)$$

N is now the number of complete cycles separating the phases of the two stations. In the two records we identify peaks of waves that correspond to the same period T_1 , and subtract their traveling times from each other to get Δt_1 (Fig. 13.8a). The integer N is selected, as in the previous case, by trial and error so that the dispersion curve is continuous and velocities are reasonable. In this way we obtain the phase velocity for periods present on the two seismograms (Fig. 13.8b). The phase velocities obtained correspond to a mean value of the velocity of the structure along the trajectory between the two stations. This method is no longer used and has been substituted by those using Fourier analysis.



Fig. 2.1

Damage of the 10 January 2010 earthquake $M_w = 7$, in Port-au-Prince, Haiti (U.S. Geological Survey).



Fig. 2.2

Surface rupture of the San Andreas fault, California. View along Carrizo Plains (U.S. Geological Survey, photo R. E. Wallace).

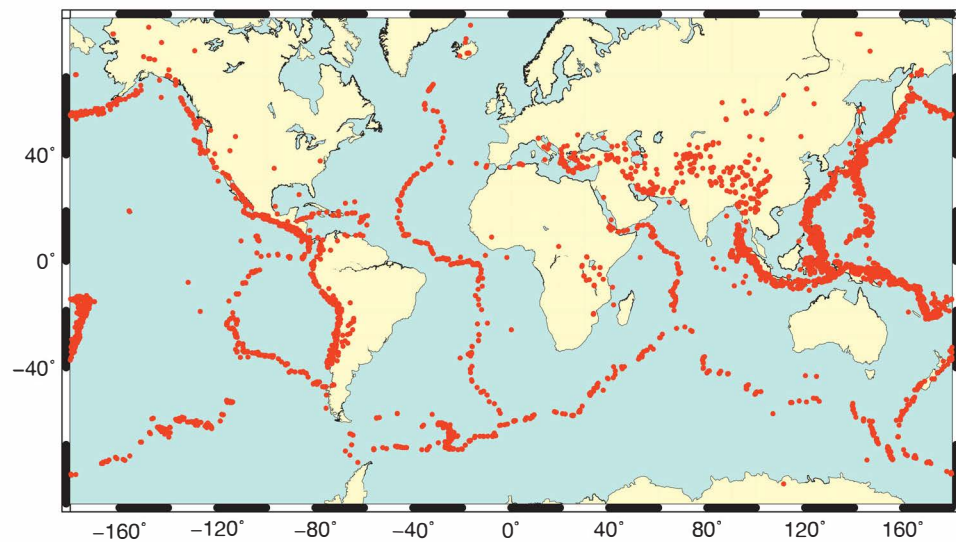


Fig. 2.3 World map of shallow earthquakes ($h < 60$ km) for the period 2000–2015, $M > 5$ (data from NEIC, U.S. Geological Survey).

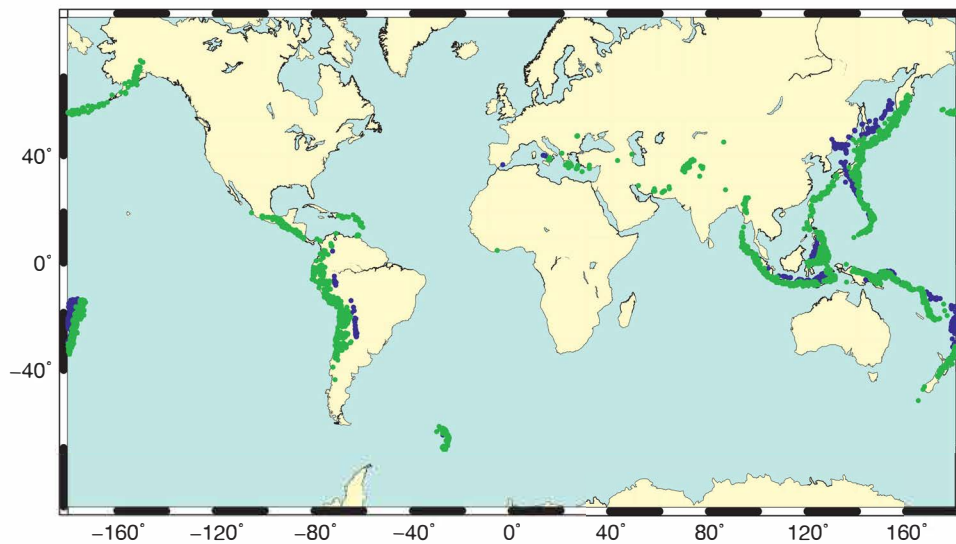


Fig. 2.4 World map of intermediate depth ($60 \text{ km} < h < 300 \text{ km}$, green) and deep ($300 < h < 700 \text{ km}$, blue), earthquakes period 2000–2015, $M > 5$ (data from NEIC, U.S. Geological Survey).

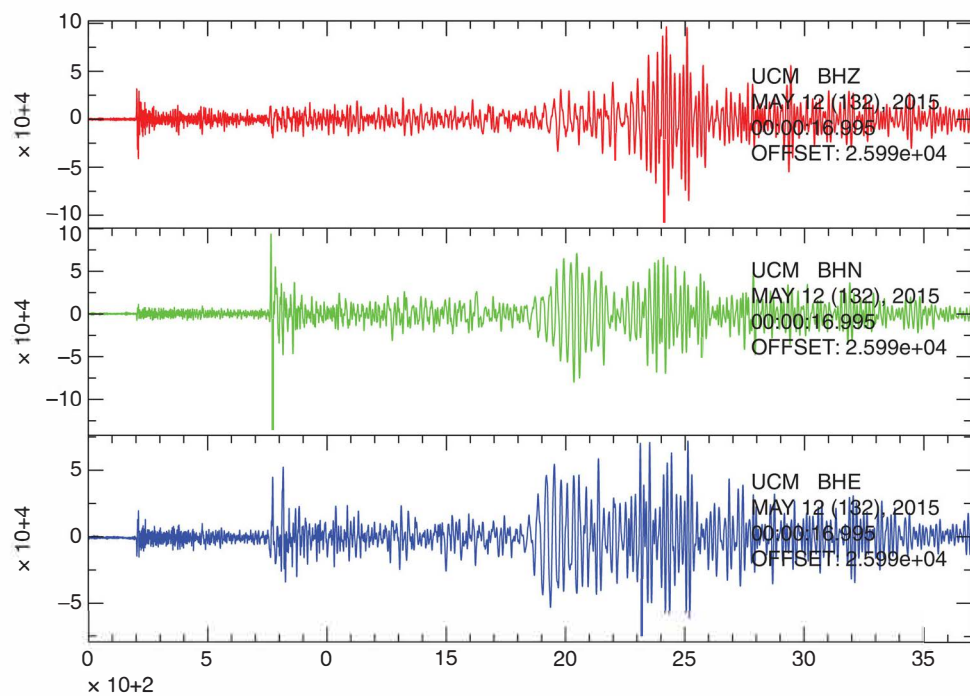


Fig. 3.9

Seismograms of the three components (BHN north–south, BHE east–west, and BHZ vertical) of the UCM station (Western Mediterranean Network, Universidad Complutense de Madrid, Orusco de Tajuña, Madrid, Spain) for the Nepal earthquake of 12 May 2015.



Fig. 3.16

Deployment of an OBS at the Alboran Sea (between Spain and Morocco) by a ship of the Spanish Navy (courtesy of Dr. A. Pazos, from the Real Observatorio de la Armada, San Fernando).

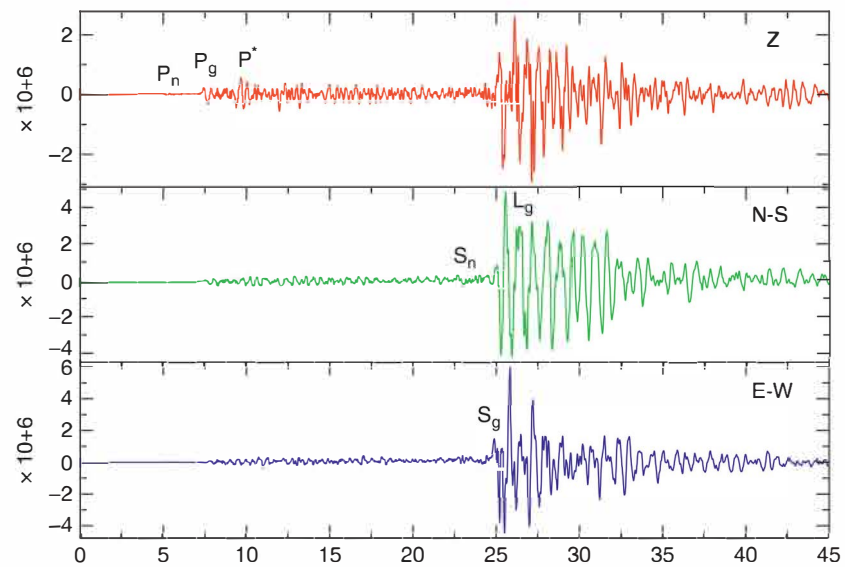


Fig. 11.10

A seismogram of a local earthquake in Ossa de Montiel, Spain (23 February 2015), recorded at the UCM BB station (Madrid, Spain) at 150 km distance, showing crustal phases.

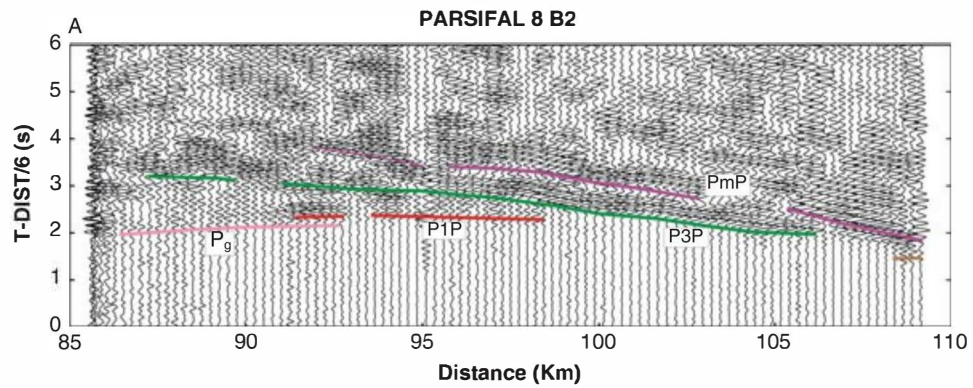


Fig. 11.11 An observed seismic record section of a refraction seismic profile offshore in the Gulf of Cadiz, Spain. P_g , P_1P , P_3P , and P_mP phases are clearly recorded; P_1P and P_3P are reflections at intermediate crustal layers (courtesy of T. Medialdea).

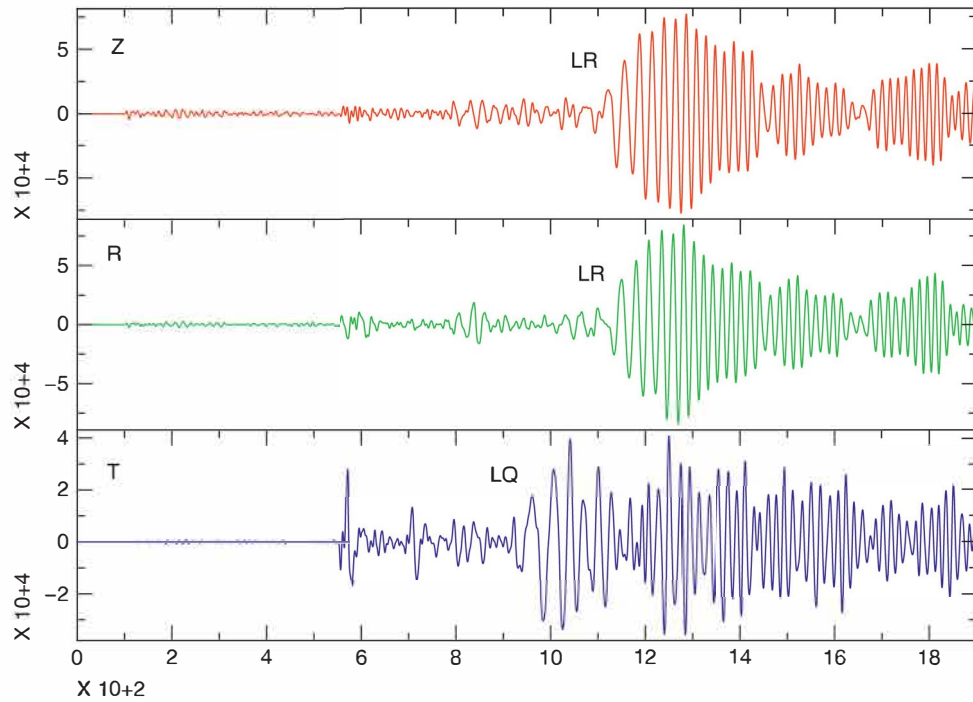


Fig. 13.10 LR and LQ waves on radial, transverse, and vertical components from the record of the broad-band station ANMO of a shallow earthquake in Andreanof Islands, Aleutian Is. (19 December 2007, $M_w = 7.2$).

Saint Vincent 12-02-2007, h=30 km

$M_0 = 0.770E+18$ Nm $M_w = 5.86$

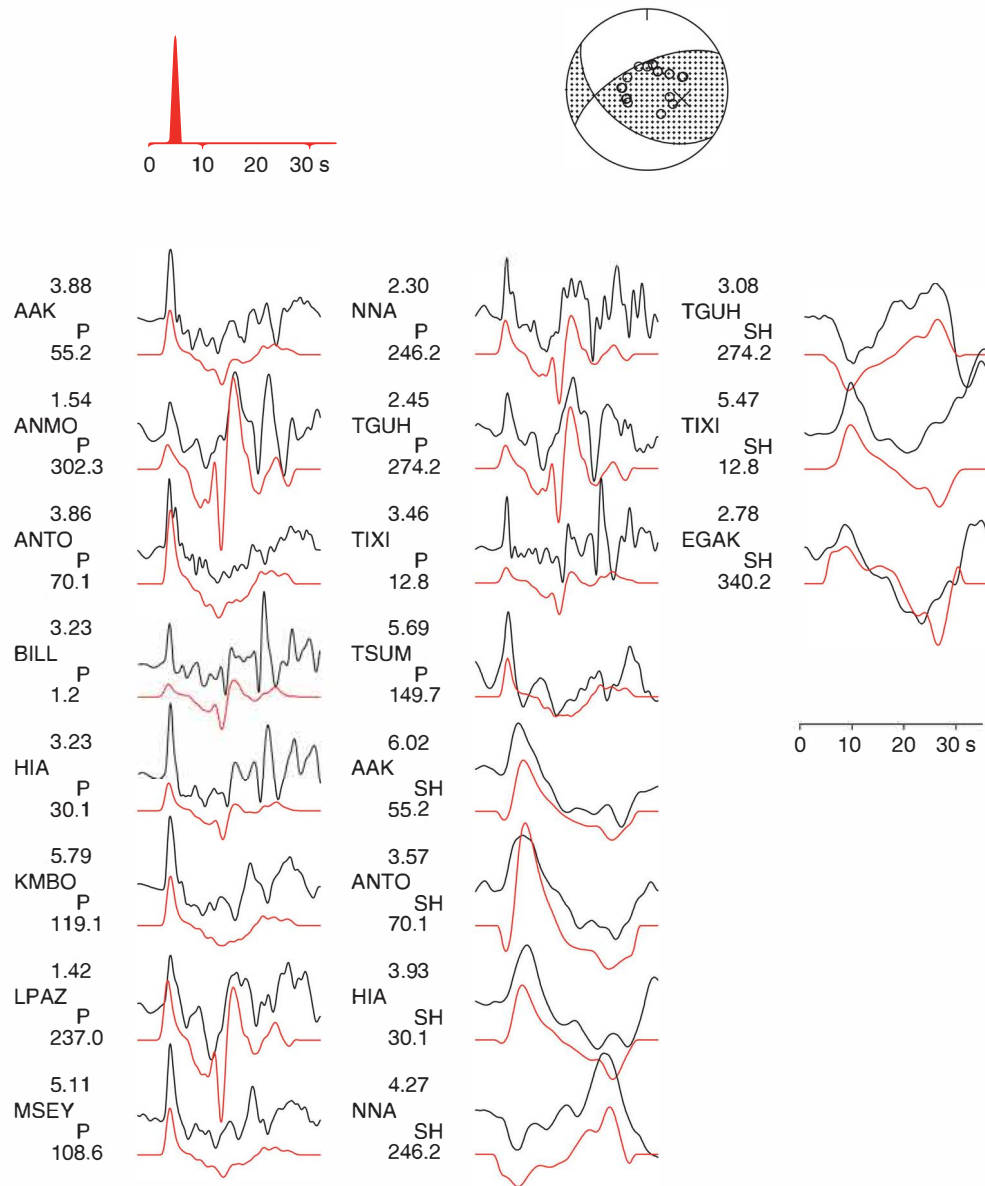


Fig. 20.8 Modeling of P and SH waves for the Saint Vincent earthquake of 12 February 2007.

Chile 02-01-2011 h=30 km

$M_0 = 0.397E+20$ Nm $M_w = 7.00$

H=30.0km var=0.21

max slip = 1.72m

(350.,14., 84.)

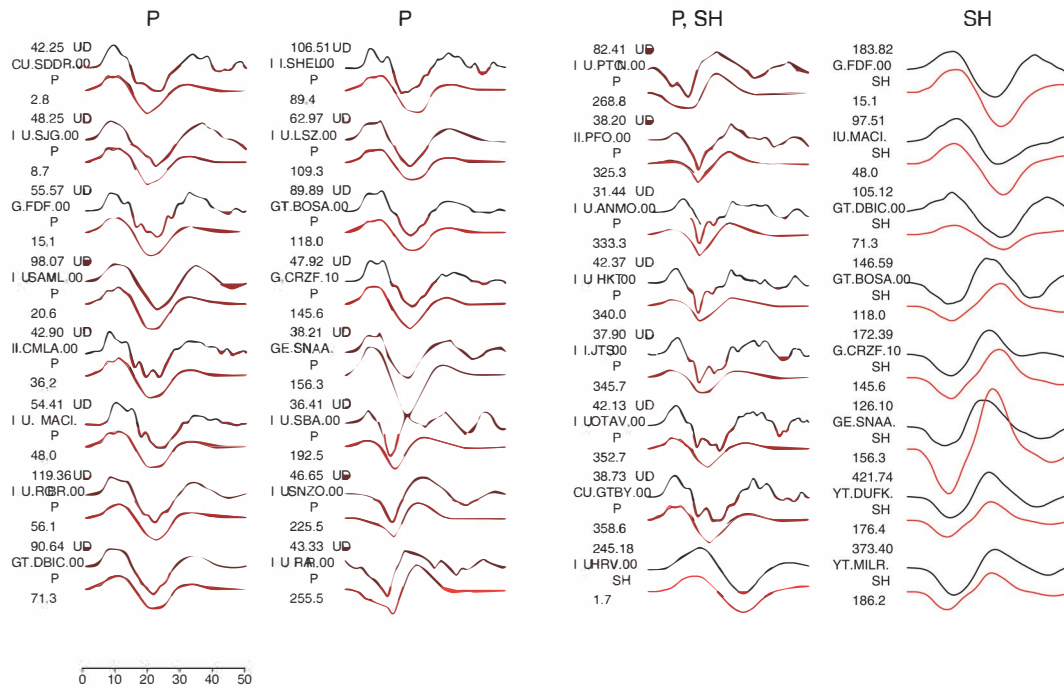
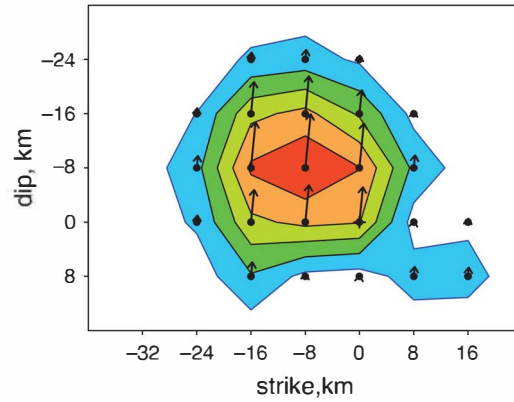
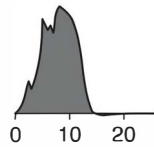
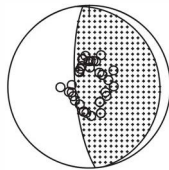


Fig. 20.11 Distribution of the slip over the focal sphere for the Chile earthquake of 2 January 2011, $M_w = 7.0$ (courtesy of C. Pro).

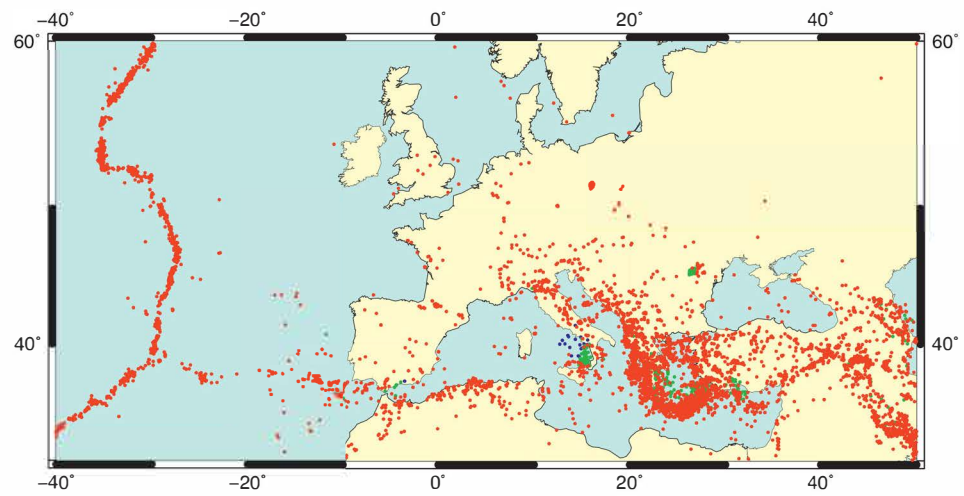


Fig. 21.1

Map of earthquakes in the Azores–Mediterranean region for the period 2000–2015 $M > 4$ (red shallow ($h < 60$ km), green intermediate depth ($60 < h < 300$ km), blue deep ($h > 300$ km)) (data from NEIC, U.S. Geological Survey).

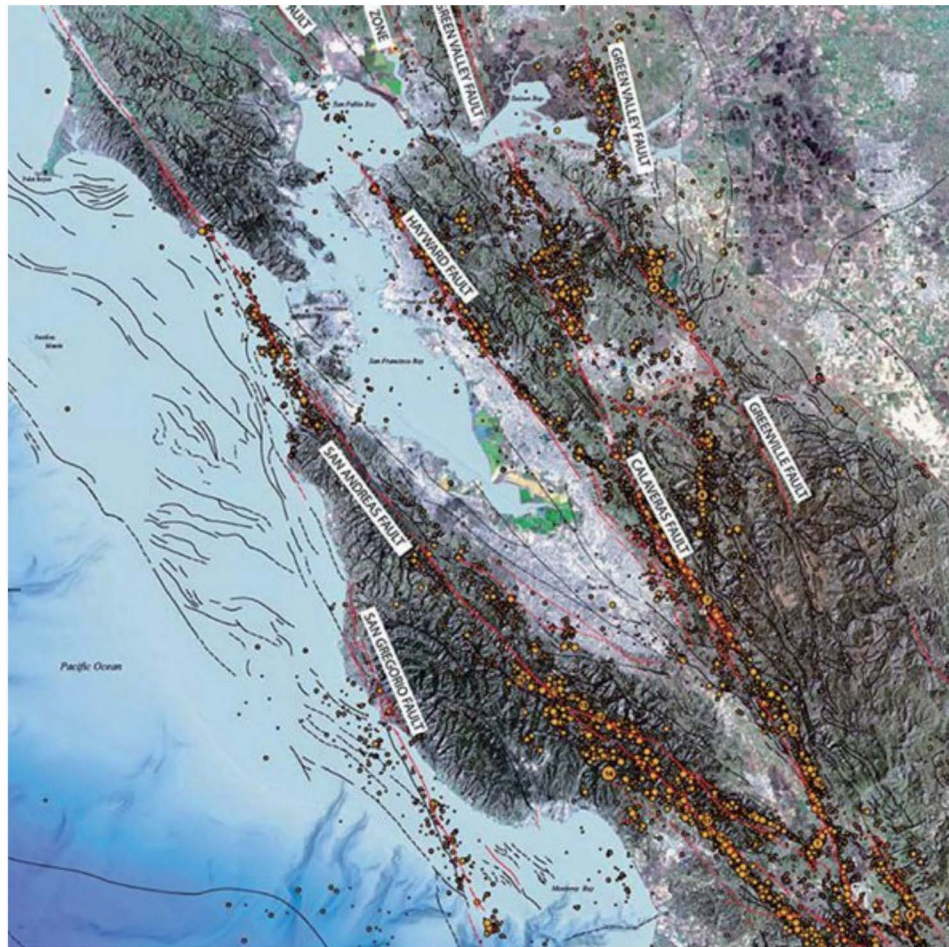


Fig. 21.2

Geological faults and earthquake distribution in California (with permission of U.S. Geological Survey).

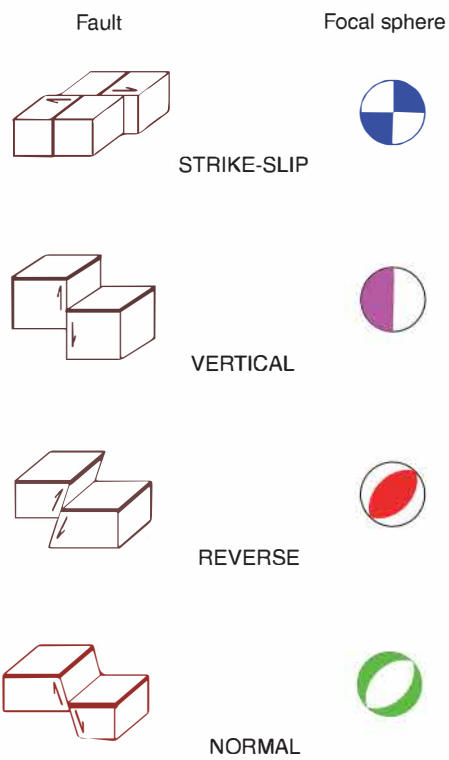


Fig. 21.7 The correspondence between faults and diagrams of the focal mechanism (horizontal projection of the focal sphere).

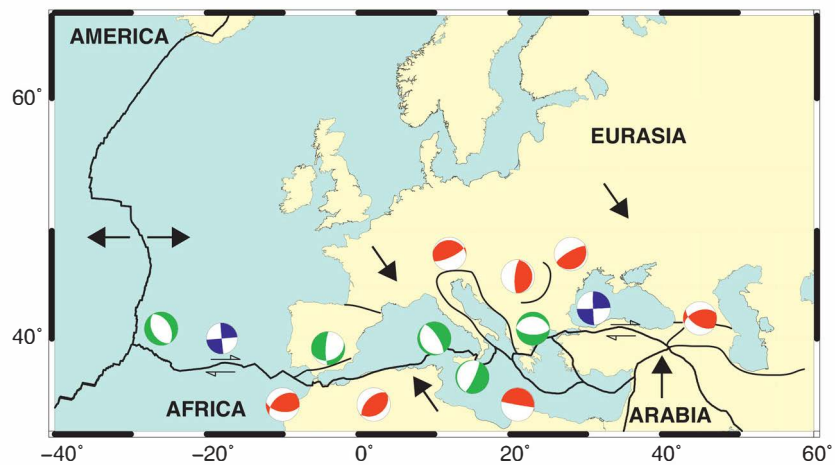


Fig. 21.9 The seismotectonic framework for the Azores-Mediterranean region.

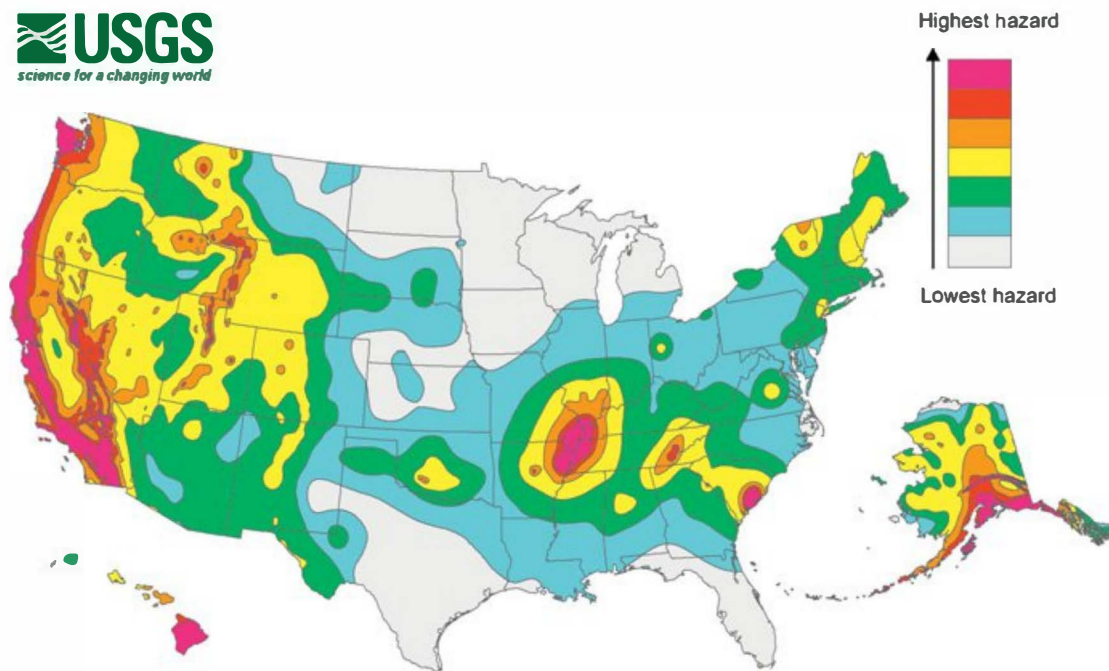


Fig. 21.11 A seismic hazard map for the USA (with permission of U.S. Geological Survey).

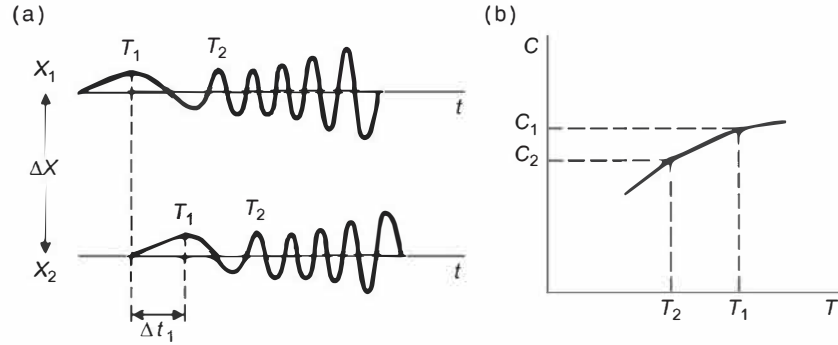


Fig. 13.8

Determination of the phase velocity using two stations. (a) Identification of the same phase for the same period at both stations and determination of the time interval. (b) The phase velocities corresponding to each period.

13.6 Determination of phase and group velocities. Fourier analysis

13.6.1 Fourier analysis of seismograms

The recording of a dispersed train of surface waves at a station located at a given distance from the epicenter is a real function of time $u(t)$. As we saw in Section 5.2.2 its Fourier transform is a complex function of frequency $F(\omega)$ (Appendix 4):

$$F(\omega) = \int_0^{\infty} u(t) e^{-i\omega t} dt, \quad (13.50)$$

where $F(\omega)$ is the complex spectrum $F(\omega) = R(\omega) + iI(\omega)$ and can also be expressed as

$$F(\omega) = A(\omega) e^{i\Phi(\omega)}. \quad (13.51)$$

$A(\omega)$ is the amplitude spectrum and $\Phi(\omega)$ is the phase spectrum. By using the inverse transform we pass from $F(\omega)$ to $u(t)$:

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega. \quad (13.52)$$

Seismic waves can be studied in the time domain $u(t)$ or in the frequency domain $F(\omega)$. Spectra show the contribution of each frequency to the waves observed. The record at a given distance of a dispersed train of surface waves, as we have already mentioned, shows at each time only the presence of waves corresponding to instantaneous frequencies. However, all frequencies are really present with more or less energy and their amplitudes and phases can be obtained by using Fourier transformation. In this way, we can calculate phase and group velocities for all frequencies, without being limited to instantaneous values like in the previous section.

In equations (13.50) and (13.52), $u(t)$ is a continuous real function of time defined from zero to infinity. In practice, it is only defined for a finite length and sampled at certain time intervals; that is, it is formed by a finite number of discrete values, u_i , for $i = 1$ to N at time intervals δt . Its Fourier transform is also discrete, F_k , for $k = 1$ to $N/2$, corresponding to frequencies between 0 and $\pi/\delta t$ at frequency intervals of $\delta\omega = 2\pi/(N\delta t)$. The highest frequency, $\omega_N = \pi/\delta t$, is the *Nyquist frequency* (named after Harry Nyquist) and depends only on the sampling interval (Appendix 4).

Since early work by Yasuo Satô (1955), many methods for determination of phase and group velocities of surface waves on the basis of Fourier analysis have been developed, especially since 1960 with the availability of digital computers and fast methods of computation of Fourier transforms (Dziewonski and Hales, 1972). We show the basic ideas of two of these methods as examples.

13.6.2 The phase velocity

Fourier analysis can be easily applied for the determination of the phase velocity from records of surface waves at two stations. The method is similar to that of instantaneous frequencies, but now we use all of the frequencies obtained from Fourier spectra. At each station the amplitude and phase spectrum, starting at times t_1 and t_2 , is determined. For a distance Δx between stations and an interval $\Delta t = t_2 - t_1$ between the times of starting analysis at the stations, the difference between the phases is

$$\Phi_2(\omega) - \Phi_1(\omega) = \delta\Phi(\omega) + 2\pi N = k\Delta x - \omega\Delta t. \quad (13.53)$$

The term $2\pi N$ is, as before, added to compensate for the number of complete cycles separating the two phases. On dividing by k and solving for the phase velocity ($c = \omega/k$), we find an equation similar to (13.49):

$$c(\omega) = \frac{\Delta x}{(1/\omega)[\delta\Phi(\omega) + 2\pi N] + \Delta t}. \quad (13.54)$$

If, for the two seismograms, we take the same starting time for calculating the Fourier transform, then $\Delta t = 0$. On expressing equation (13.54) as a function of the period, we obtain

$$c(T) = \frac{\Delta x}{T[\delta\Phi(T) + N]}. \quad (12.55)$$

Equations (12.54) and (12.55) give the phase velocity for all frequencies obtained in the Fourier transform. In practice, since the seismogram is a sampled function of limited length, discrete Fourier transformation provides only a limited number of frequencies (Appendix 4). These limitations are present in all methods based on Fourier analysis of sampled functions of finite length.

13.6.3 The group velocity

Most methods used to obtain the group velocity using Fourier analysis are based on the application of a band-pass filter called a *multiple filter* that is centered on a frequency that

takes various values (Landisman *et al.*, 1969). We will present only the basic ideas of this method. We define a Gaussian filter in the frequency domain centered on a certain frequency ω_n and with a width related to the constant α , in the form

$$H(\omega_n, \omega) = e^{-\alpha(\omega - \omega_n)^2 / \omega_n^2}. \quad (13.56)$$

The Gaussian function has the advantage that its transform is also a Gaussian function. The frequency ω_n takes successive values for which we want to calculate the group velocity. The Gaussian filter (13.56) is applied to the Fourier transform of the dispersed surface waves $u(t)$. The result, after eliminating the effects of initial and instrumental phases, can be written as

$$h(\omega_n, t) = \int_{-\infty}^{\infty} A(\omega) e^{-\alpha(\omega - \omega_n)^2 / \omega_n^2} \cos(kx - \omega t) d\omega, \quad (13.57)$$

where $h(\omega_n, t)$ is the inverse transform of the filtered function. Because of the shape of the filter, $A(\omega)$ and $k(\omega)$ can be expressed using a Taylor expansion about the central frequency ω_n :

$$A(\omega) = A(\omega_n) + (\omega - \omega_n) \frac{dA}{d\omega} = A_n + (\omega - \omega_n) A'_n, \quad (13.58)$$

$$k(\omega) = k(\omega_n) + (\omega - \omega_n) \frac{dk}{d\omega} = k_n + (\omega - \omega_n) k'_n. \quad (13.59)$$

On substituting these into (13.57), we obtain

$$h(\omega_n, t) = g(\omega_n, t) \cos(k_n x - \omega_n t + \varepsilon_n) \exp[-\omega_n^2 (k'_n x - t)^2 / (4\alpha)], \quad (13.60)$$

where

$$g(\omega_n, t) = (\pi/2)^{1/2} \omega_n \{A_n^2 + [A'_n \omega_n (k'_n x - t)]^2 / (4\alpha^2)\}^2, \quad (13.61)$$

$$\varepsilon_n(\omega_n, t) = \tan^{-1} [A'_n \omega_n (k'_n x - t) / (2\alpha A_n)]. \quad (13.62)$$

We can define the *instantaneous amplitude* $a(t)$ and *phase* $\Phi(t)$ for an arbitrary function $f(t)$ (Dziewonski and Hales, 1972) in the form,

$$a(t) = [f(t)^2 + q(t)^2]^{1/2}, \quad (13.63)$$

$$\Phi(t) = \tan^{-1} [q(t)/f(t)], \quad (13.64)$$

where $q(t)$ is the inverse transform of the Fourier transform of $f(t)$ after introducing a phase shift of $\pi/2$, or, equivalently, after interchanging its real and imaginary parts. The function $a(t)$ can be considered as the envelope of $f(t)$. On applying this operation to $h(\omega_n, t)$, we obtain

$$a_n(t) = g(\omega_n, t) \exp[-\omega_n^2 (k'_n x - t)^2 / (4\alpha)], \quad (13.65)$$

$$\Phi_n(t) = \omega_n t - k_n x - \varepsilon_n. \quad (13.66)$$

From the principle of a stationary phase it follows that $d\Phi/dt = -\omega$ (13.37), where ω is the instantaneous frequency corresponding to each time. Maxima of $a(t)$ correspond to times for which $da/dt = 0$. Taking the derivative in (13.65) and putting it equal to zero, this is satisfied for $t_n = k'_n x$. However, $k'_n = dk_n/d\omega = U_n$ is precisely the group velocity associated with the maxima of $a_n(t)$, which is given by

$$U_n = x/t_n, \quad (13.67)$$

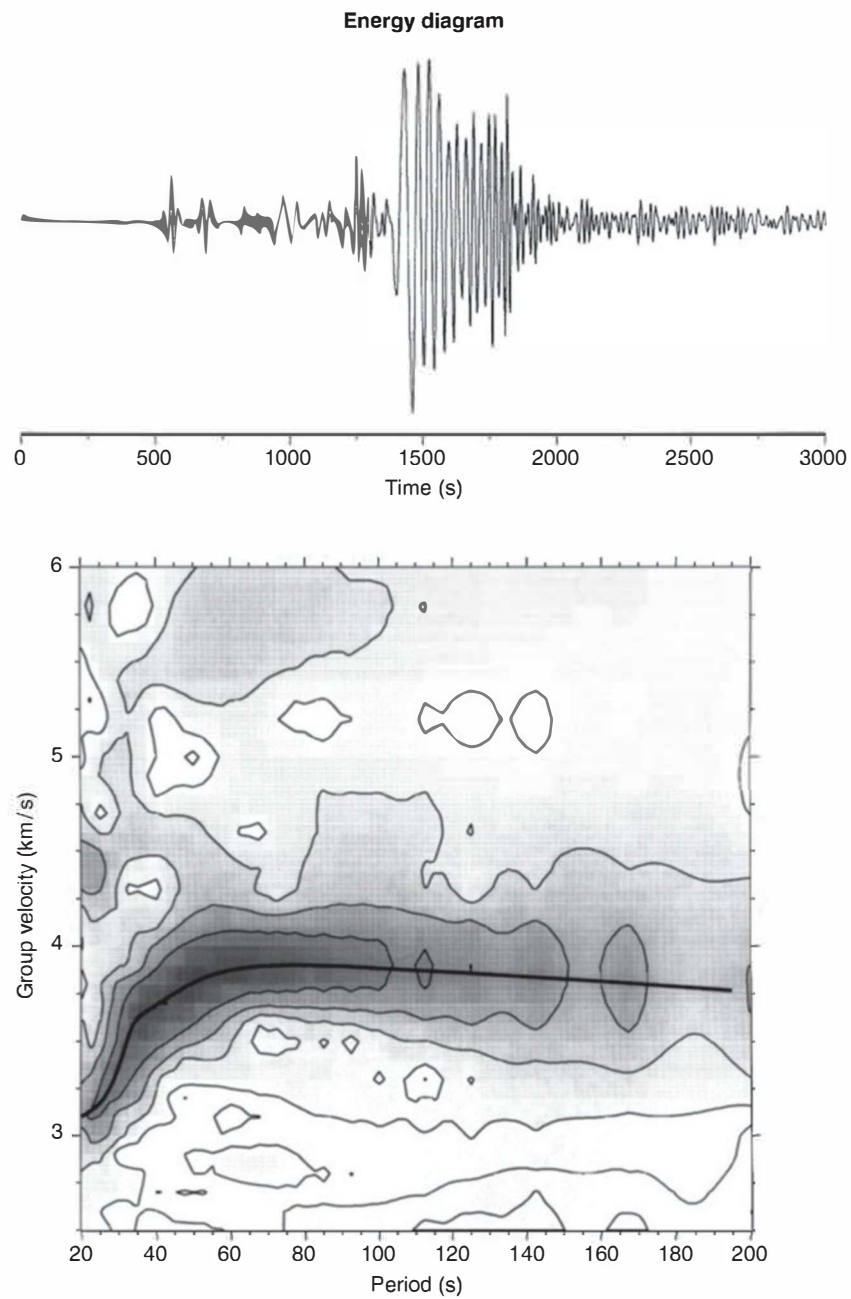
where x is the distance from the station to the epicenter and t_n is the time corresponding to the arrival of the maximum of $a_n(t)$. Since this function is the envelope of $h(\omega_n, t)$, that is, of the original function $u(t)$ filtered by the Gaussian function centered at the frequency ω_n , this maximum represents the energy contained in $u(t)$ corresponding to the frequency ω_n . In consequence, t_n is the group travel time associated with the frequency ω_n . By giving various values to ω_n , we find their corresponding group travel times t_n and, according to (13.67), we find the values of their group velocities. The method is actually equivalent to the method of measuring the times peaks and troughs on a seismogram, but it is not restricted to instantaneous frequencies. If, for the same frequency ω_n , there is more than one maximum of $a_n(t)$, they correspond to different modes of surface waves. The greatest travel times (lowest velocities) correspond to the fundamental mode.

Since, in practice, we always use a sampled function $u_m (m = 1, \dots, M)$, the function a_{mn} has a matrix form and is called the *energy diagram*, where the subindex m refers to times and n refers to frequencies used in the filter. In this matrix, for each value of the frequency ω_n , we select the time t_n that corresponds to the maximum of a_{mn} . This time is the group travel time associated with the maximum of energy corresponding to the frequency ω_n . By dividing the epicentral distance by t_n , we find the group velocity U_n (13.67) corresponding to each frequency ω_n or period T_n (Fig. 13.9). This type of analysis can also be used to filter the signal with a selected distribution of group velocities. In this form, we can separate the energies propagated by the various modes, for example, that of the fundamental mode from those of higher modes. Filters of this type are called *group-velocity filters*.

13.7 Dispersion curves and the Earth's structure

13.7.1 Observations

Surface waves in the Earth are observed on seismograms of distant surface earthquakes as long trains of dispersed waves with large amplitudes (Fig. 13.10). Dispersion is easily detected, first arrivals corresponding to waves of longer periods. As has been mentioned, the periods present on seismograms correspond to instantaneous frequencies. As we saw in Chapter 12, Love waves are registered only in the horizontal components whereas Rayleigh waves, which are polarized in the vertical plane, are registered both in horizontal and in vertical components. If we rotate the two horizontal components to make them coincide with the radial and transverse directions with respect to the orientation from the station to the epicenter, Love waves (LQ) are

**Fig. 13.9**

Determination of the group velocity using the multiple filter method. Dispersed surface waves and the energy diagram with group velocities and periods are shown for the Alaska earthquake of 20 November 1993, recorded at the HRV station, $\Delta = 50^\circ$ (courtesy of L. A. Rivera).

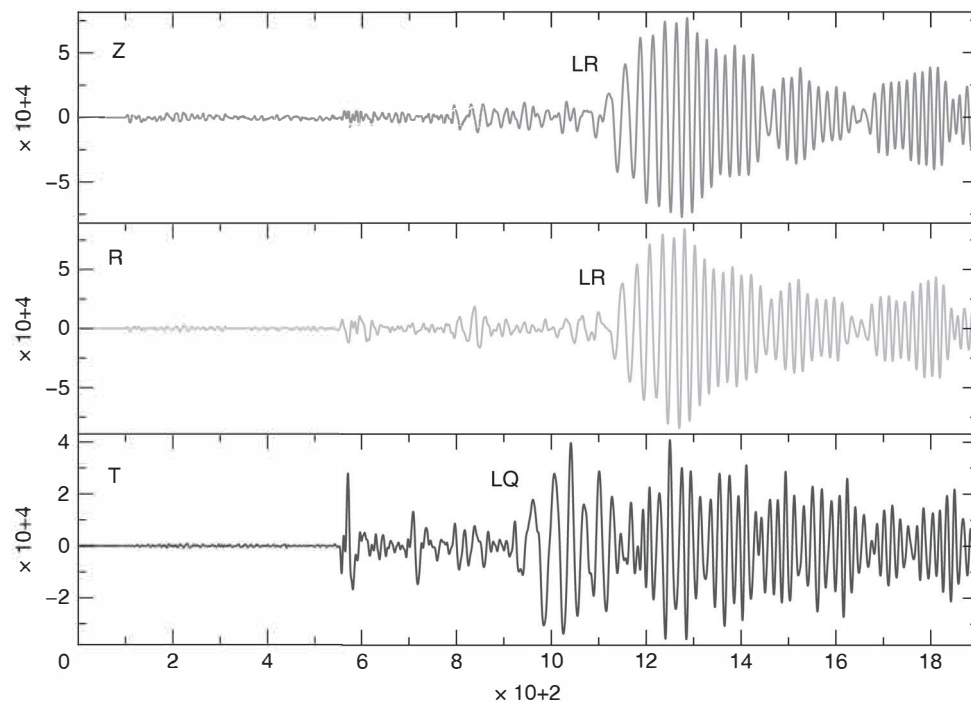


Fig. 13.10

LR and LQ waves on radial, transverse, and vertical components from the record of the broad-band station ANMO of a shallow earthquake in Andreanof Islands, Aleutian Is. (19 December 2007, $M_w = 7.2$). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

recorded only in the transversal component and Rayleigh waves (LR) are recorded only in the radial (Fig. 13.10). Love waves of long periods (60–300 s) are also called G waves (in honor of Gutenberg). For these periods, the dispersion curve is practically flat with a velocity of about 4.4 km s^{-1} and waves have an almost impulsive form.

Love and Rayleigh waves of short periods (8–12 s) in continental trajectories are channeled in the upper crust and are known as Lg and Rg waves. For periods between 60 s and 300 s, Love and Rayleigh waves travel mainly through the mantle and are called *mantle waves*. For large earthquakes, surface waves that travel around the Earth more than once are observed. These waves are designated with a subindex, G_1 , G_2 , G_3 , etc. for Love waves, and R_1 , R_2 , R_3 , etc. for Rayleigh waves (Fig. 13.11). Groups G_1 and R_1 are direct waves from the epicenter to the station and G_2 and R_2 are waves that arrive at the station traveling in the opposite direction. For higher subindexes, waves with an odd subindex circle the Earth, leaving the epicenter in the direction of the station; and those with an even subindex do this leaving in the opposite direction.

13.7.2 Interpretation

The forms of dispersion curves of phase and group velocities of Love and Rayleigh waves depend on the characteristics of the medium (Section 12.3). Thus, their study

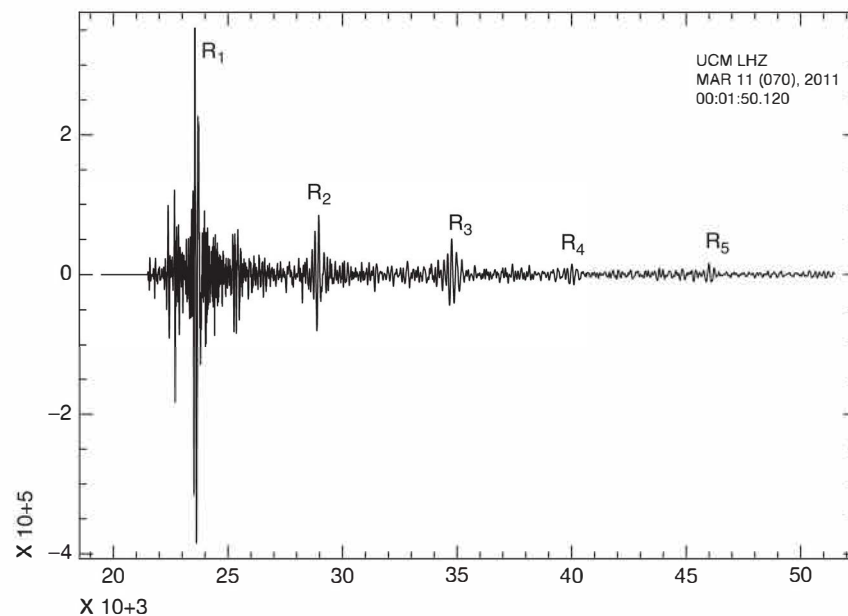


Fig. 13.11

Records of several groups of Rayleigh waves that circle the Earth. They were recorded at the broad-band station UCM (Madrid) from the earthquake of Tohoku, Japan (11 March 2011); $M = 9.0$.

allows determination of the structure present along their trajectories. Since surface waves travel along the surface of the Earth, their penetration into its interior depends on their wave length. For periods less than 60 s they are affected by the crust and for periods between 60 and 300 s they are affected by the mantle. They provide average values of the structure along their trajectory and do not give details of lateral variations. However, the analysis of surface waves along many trajectories permits separation between the effects of the various structures they have crossed. With this type of analysis, called *regionalization*, we can separate structures for different regions, such as oceanic regions with various ages of sea floor, shields, orogenic regions, rifts, etc. Owing to the difficulty of interpreting surface waves of short periods (less than 10 s) because of the presence of large lateral heterogeneities, in general, regionalization does not give good results for small regions of the crust.

The first studies of the structure of the Earth's crust and mantle by means of the analysis of Rayleigh and Love waves were those by Ewing and Dean S. Carder in the 1930s. Later, during the 1950s, we have the work of Satô, Press, Jack Oliver, and James T. Wilson (Press, 1956; Oliver, 1962) among others. The rapid development of both long period seismographic instrumentation and digital computers during the 1960s gave a great impetus to these studies. During the 1960s and 1970s many studies of crust and mantle structures along a large number of trajectories, crossing all types of structure, oceanic, continental, shields, rifts, etc., were completed. Computers also made possible the calculation of theoretical dispersion curves for stratified media with many layers, needed for the interpretation of observations. Observed dispersion curves were at first inverted by comparison of observed

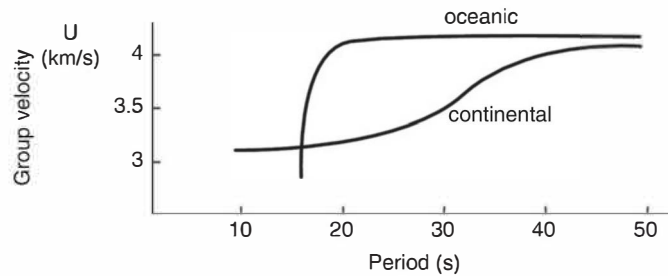


Fig. 13.12 Theoretical group velocity dispersion curves of Rayleigh waves for continental and oceanic structures. Continental crust, $H = 30$ km, $\beta = 3.5$ km s⁻¹, upper mantle, $\beta = 4.6$ km s⁻¹, oceanic crust $H = 10$ km, $\beta = 3.2$ km s⁻¹, and upper mantle, $\beta = 4.3$ km s⁻¹.

and theoretical curves for various types of models. More recently, observed data have been inverted directly, including compensation for errors in observations and resolution of models.

One of the first results from the analysis of dispersion curves of surface waves was the difference between crust and upper mantle structures under oceans and those under continents. Purely oceanic trajectories show that group velocities of Rayleigh waves with periods 20–100 s have a very constant value of about 4 km s⁻¹ and that those with shorter periods, 10–20 s, have a sharp fall due to the water layer. For continental trajectories for periods shorter than 60 s, the group velocity decreases from 4 km s⁻¹, and for 10 s it reaches a minimum of about 3 km s⁻¹. This is due to the thickness of the crust being greater and to the presence of sediments with low velocities. For periods greater than 150 s, dispersion curves for continents and oceans are similar due to the influence of deeper parts of the mantle. These general characteristics can be explained by invoking the very simple model of oceanic and continental crusts, and the upper mantle shown in Fig. 13.12. A summary of observed dispersion curves for group velocities of Rayleigh and Love waves for a large range of periods is shown in Fig. 13.13. Along oceanic and continental trajectories, Love waves have higher velocities than Rayleigh waves. For oceanic paths, Love waves have a practically constant velocity of about 4.4 km s⁻¹ for periods in the range 20–400 s. For periods shorter than about 80 s, Love and Rayleigh waves for continental paths have lower values than for oceanic paths. There is a sharp drop in velocity for both Rayleigh and Love waves of about a 10–20 s period for oceanic paths. For Rayleigh waves there is minimum for a period of about 200 s that is produced by the influence of the low-velocity layer in the mantle. Characteristics of the crust and mantle found from the analysis of surface waves agree with those found from travel times of body waves (Sections 11.3 and 11.4).

A comparison among dispersion curves for various types of path reveals that their structures are different (Knopoff, 1972). For Rayleigh waves and periods in the range 60–140 s, phase velocities are highest for continental shields and lowest for rift zones (Fig. 13.14). These curves reveal the effect of the differing thickness of the lithosphere and the influence of the asthenosphere. The thickness of the lithosphere is greater for

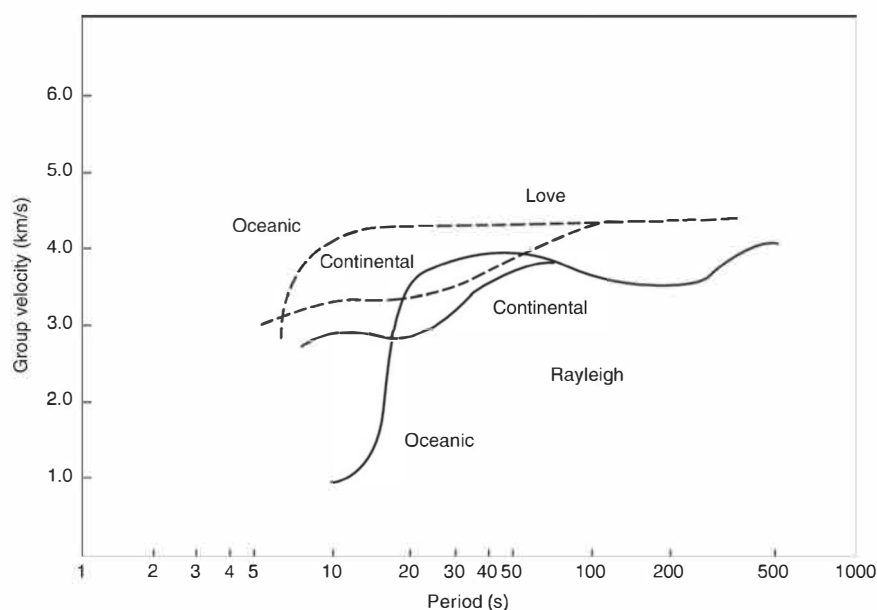


Fig. 13.13 Group velocity dispersion curves for Rayleigh and Love waves (modified from Bullen and Bolt (1985)).

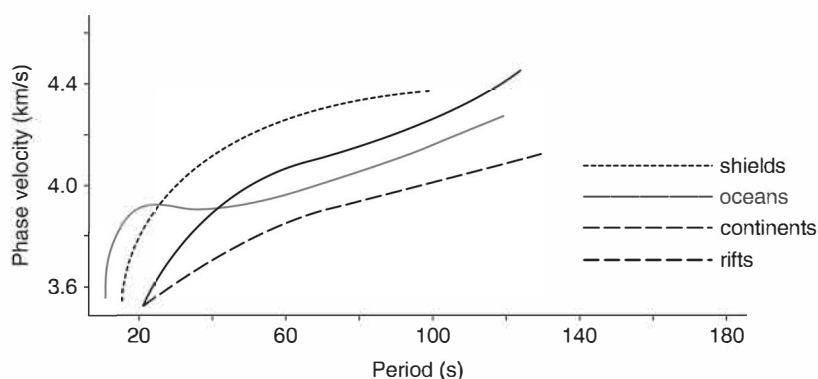


Fig. 13.14 Average phase velocity dispersion curves of Rayleigh waves for regions of shields, oceans, continents, and rift regions (modified from Knopoff (1972)) (with permission from Elsevier Science).

continental shields and smaller in rift zones with a shallower asthenosphere. The mean curves given in Fig. 13.14 show only the differences among very broad types of crust.

Since 1980, with the installation of global networks of seismographic stations with digital broad-band instrumentation (Chapter 3), it has been possible to analyze surface waves for various types of trajectories using not only dispersion curves but also amplitudes. These studies allow a more detailed regionalization of the structure of the crust and mantle, revealing lateral inhomogeneities through the whole mantle. Generally, these studies concern S wave velocity distributions obtained using spherical harmonics (Appendix 3).

The best results have been found for harmonics of orders between four and ten, applied to the resolution of mantle structure 300–1000 km deep (Dziewonski and Woodhouse, 1987). In this way, global three-dimensional models of the Earth's mantle have been obtained by the application of tomographic techniques to observation of surface waves. The ambient noise tomography methods are also widely used to determine crustal and upper mantle structures at regional and global scale. They are based on cross-correlations of background noise recorded at station pairs (Shapiro *et al.*, 2005)

13.8 Summary

Surface waves are dispersed, that is, their velocity depends on frequency so that they form groups of waves that travel with a group velocity different from the phase velocity of their components. The principle of a stationary phase states that for a given distance and time, energy is contained in waves with a frequency that corresponds to stationary values of the phase. This gives us the relation between phase and group velocity. Highest amplitudes form the Airy phase corresponding to the frequency for which $dU/dk = 0$. Observed group and phase velocities of dispersed surface waves can be determined directly from the instantaneous frequencies observed in seismograms. Using Fourier analysis we can extend the analysis to all frequencies. Thus, phase velocities along trajectories between two stations can be obtained from the difference between the observed phases for the same frequency. Group velocities for the path from the epicenter to a station can be obtained by applying a multiple filter to an observed train of dispersed waves. Group velocity corresponds to the quotient of the distance to the times at which the maxima of the filtered signal arrive.

Observations of Rayleigh and Love wave dispersion curves allow for determination of the average structure along their trajectories. Group velocities of Rayleigh waves with periods between 10 and 200 s show the different structures between oceanic and continental crust and upper mantle. Detailed studies of phase velocity dispersion curves for different types of trajectory show the different structures of the crust and upper mantle in regions of oceans, shields, continents, and rifts.

13.9 Problems

- P13.1 For the third higher mode of Love waves in a layer of thickness H , velocity $\beta/2$, and rigidity $\frac{3}{4}\mu$ over a half-space of velocity β and rigidity μ , find:
- the maximum wave length;
 - the depths at which the displacements are null;
 - the wave length for which the phase velocity $c = \frac{3}{4}\beta$.
 - the dispersion curve c/β versus H/λ .

- P13.2 For the hypothetical case of a layer of thickness H and velocity β' over a half-space of velocity β , if the phase changes on the free surface and at the plane of contact of the layer and the medium are $\pi/4$, and

$$-\tan^{-1} \left(\frac{(c^2/\beta'^2 - 1)^{1/2}}{(1 - c^2/\beta'^2)^{1/2}} \right),$$

respectively,

- (a) Write the dispersion equation using the condition of constructive interference.
 - (b) Find the cut-off frequencies of the fundamental and first higher mode.
 - (c) Draw the dispersion curves for the fundamental and the first higher mode using c/β and H/λ for $\beta = 2\beta'$.
- P13.3 On a seismogram of dispersed Rayleigh waves given in the EMP13_3, the origin time of the earthquake is 09h30m28s and the epicentral distance 6715 km. Using the method of peaks and troughs (instantaneous frequencies):
- (a) Identify the arrival times for the periods.
 - (b) Calculate the group velocity for each period.
 - (c) Draw the dispersion curve of the group velocity U versus the period.
 - (d) Calculate the amplitude of the Airy phase (1 count = 10^{-9} ms $^{-1}$).
- P13.4 Use the EMP13_3 and EMP13_4 corresponding to the same earthquake as that of [Problem 13.3](#). The EMP13_4 corresponds to a second station on a great circle path at a distance of 366 km from the first. Using the method of peaks and troughs (instantaneous frequencies):
- (a) Calculate the phase velocity for each period.
 - (b) Draw the dispersion curve of the phase velocity c versus the period.

Up to this chapter we have considered motion generated by earthquakes in the Earth as waves propagating inside an unbounded infinite elastic medium (P and S body waves) or along an infinite free surface and layered medium (surface waves). The finite dimensions of the Earth have not been considered. If the wave lengths of motion are very small compared with the dimensions of the Earth this is a good approximation. For example, body waves have relatively high frequencies corresponding to wave lengths of between about 5 km and 150 km, which are small compared with the Earth's radius (6370 km). But surface waves may have wave lengths of the order of the Earth's radius. For those waves the Earth cannot be considered as an infinite medium and its finite dimensions must be taken into account. For this reason, we have to consider the complete problem of motion of an elastic body of finite dimensions with no conditions imposed on its frequency, which leads into *normal mode* theory. In this chapter we will give the fundamentals of this theory starting with the simple problems of the vibration of a finite elastic string and a rod and proceed to the vibrations of an elastic sphere as an approximation for the Earth. These results will be applied to the study of the Earth's *free oscillations* and their observations following large earthquakes, such as the 2011 Japan earthquake. From this point of view the Earth reacts to an earthquake by vibrating as a whole, in the same way as a bell when it is hit. Elastic displacements generated by earthquakes can, thus, be obtained as the sum of normal modes. This is an alternative approach to the problem of wave propagation in the Earth that takes into account its finite dimensions.

14.1 Standing waves and modes of vibration

As an introduction, let us consider the phenomenon of *standing waves* formed by the sum of waves propagating in opposite directions. For the one-dimensional case, let us assume waves with the same frequency and amplitude that propagate in both directions. For sinusoidal waves, according to (5.19),

$$u(x, t) = A[\sin(kx + \omega t + \phi_1) + \sin(kx - \omega t + \phi_2)], \quad (14.1)$$

where ϕ_1 and ϕ_2 are the initial phases. Applying the relation $\sin(a + b) + \sin(a - b) = 2 \sin a \cos b$, we can write (14.1) in the form,

$$u(x, t) = 2A \sin(kx + \phi'_1) \cos(\omega t + \phi'_2), \quad (14.2)$$

where $\phi'_1 = \phi_1/2 + \phi_2/2$ and $\phi'_2 = \phi_1/2 - \phi_2/2$. Since in (14.2) we do not have the propagating term $k(x \pm ct)$, this expression corresponds to standing waves such that the dependences of x and t are separated. For each value of t , $u(x)$ is a sine function of x with wave length $\lambda = 2\pi/k$ and, for each value of x , $u(t)$ is a cosine function of t with period $T = 2\pi/\omega$.

If we impose the condition that, for $x = 0$, the amplitude of the standing wave is zero for all values of t , we obtain

$$u(x, t) = 2A \sin(kx) \cos(\omega t + \phi), \quad (14.3)$$

where $\phi = \phi_1$. If we further impose that, for another value, $x = L$, the amplitude is also zero for all t , then we have

$$2A \sin(kL) = 0 \quad \text{and} \quad kL = (n + 1)\pi, \quad n = 0, 1, 2, 3, \dots \quad (14.4)$$

In this way we have introduced a finite dimension into the problem, the distance L . Hence, in order to fulfill both conditions, the wave number k must have certain discrete values, namely,

$$k_n = \frac{(n + 1)\pi}{L}; \quad n = 0, 1, 2, 3, \dots \quad (14.5)$$

Since the velocity of the waves that travel in opposite directions has the same constant value c (14.1), frequencies of standing waves that satisfy the two imposed conditions are also limited to certain discrete values,

$$\omega_n = \frac{(n + 1)\pi c}{L}; \quad n = 0, 1, 2, 3, \dots \quad (14.6)$$

Finally, standing waves that satisfy both conditions are given by

$$u_n(x, t) = 2A \sin\left(\frac{(n + 1)\pi x}{L}\right) \cos(\omega_n t + \phi); \quad n = 0, 1, 2, \dots \quad (14.7)$$

Since, for each value of n equation (14.7) is a solution, the general solution of the problem is given by their sum:

$$u(x, t) = \sum_{n=0}^{\infty} 2A \sin\left(\frac{(n + 1)\pi x}{L}\right) \cos(\omega_n t + \phi). \quad (14.8)$$

Standing waves generated by waves traveling in opposite directions with the same velocity, when they are forced to have certain values at two points, that is, when a finite dimension (in our case L) is introduced into the problem, their frequencies and wave numbers are limited to multiples of the inverse of the distance L between the two points (14.6 and 14.5). Each solution in (14.7) is called a *mode of vibration*, the lowest, that for $n = 0$, is called the *fundamental mode* and the rest are *higher modes*, *harmonics*, or *overtones*. This is the simplest case of normal mode theory. We have found this phenomenon already in the propagation of guided waves on a liquid layer (Section 12.2) and of Love and Rayleigh waves on a layer over a half-space (Sections 12.3 and 12.4). In all cases, modes are present

because of the presence in the problem of a finite dimension (for example, the layer thickness) and its relation with the wave lengths of the motion. In the problem of standing waves the largest wave length and period corresponds to the fundamental mode, $\lambda_0 = 2L$ and $T_0 = 2L/c$. The period of the fundamental mode corresponds to the time it takes for progressive waves to travel in both directions between 0 and L . For higher modes, wave lengths and periods are fractions of those of the fundamental mode ($\lambda_n = 2L/(n+1)$ and $T_n = 2L/(n+1)c$). The total motion of standing waves is the sum of all modes (14.8). These general properties are present in all problems where we have a finite dimension.

14.2 Vibrations of an elastic string of finite length

A simple case of a mechanical problem concerning vibrations of an elastic medium with finite dimensions which leads to a normal modes solution is that of the vibrations of an elastic string of length L fixed at both ends (Fig. 14.1). The transverse motion of each point of the string $y(x, t)$ depends only on the tension T in the direction of the string that is supposed to be constant. If ρ is the mass per unit length, the equation of transverse motion for a point of the string is

$$\sum_y T_y = \rho dx \frac{\partial^2 y}{\partial t^2}. \quad (14.9)$$

The forces T_y , the components of T in the y direction at a point x and at another $x + dx$, are given by $T_y = -T \sin \theta$, and $T'_y = T \sin \theta'$ (Fig. 14.2). For small angles we can approximate sines by tangents and equation (14.9) becomes

$$T(\tan \theta' - \tan \theta) = \rho dx \frac{\partial^2 y}{\partial t^2}. \quad (14.10)$$

However, $\tan \theta$ and $\tan \theta'$ are the slopes of the curve formed by the string at the two points, and, therefore, for point x , $\tan \theta = \partial y / \partial x$. For the point $x + dx$, the slope has changed and is given by

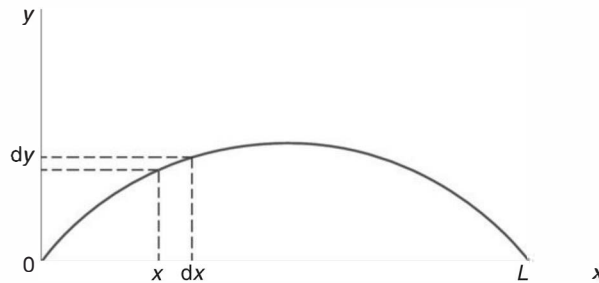


Fig. 14.1

A vibrating elastic string with both ends fixed.

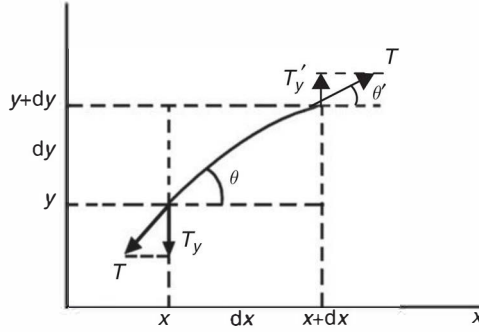


Fig. 14.2

Tension components T_y and T_y' acting on an element of an elastic string.

$$\tan \theta' = \frac{\partial y}{\partial x} + \frac{\partial}{\partial x} \left(\frac{\partial y}{\partial x} \right) dx. \quad (14.11)$$

By substitution into (14.10), we obtain

$$\frac{\partial^2 y}{\partial x^2} = \frac{\rho}{T} \frac{\partial^2 y}{\partial t^2}. \quad (14.12)$$

This equation has the same form as the wave equation (5.11) for a displacement y that propagates in the x direction. To solve equation (14.12), we apply the method of separation of variables, as in Section 5.2. On making the substitutions $y(x, t) = f(x)g(t)$ and $a^2 = T/\rho$, we obtain the following equations:

$$\frac{d^2 f}{dx^2} + \frac{\omega^2}{a^2} f = 0, \quad (14.13)$$

$$\frac{d^2 g}{dt^2} + \omega^2 g = 0, \quad (14.14)$$

where $-\omega^2$ is the constant of separation of variables. The solution of (14.12) is the product of the solutions of (14.13) and (14.14) and can be written as

$$y(x, t) = \left[A \cos\left(\frac{\omega x}{a}\right) + B \sin\left(\frac{\omega x}{a}\right) \right] [C \cos(\omega t) + D \sin(\omega t)]. \quad (14.15)$$

This equation is equivalent to (14.2), where ω is the frequency of transverse displacements of the string. The boundary conditions at both ends of the string are that they must remain fixed for all times, that is, $y(0, t) = 0$ and $y(L, t) = 0$. The first condition results in $A = 0$ and from the second,

$$y(L, t) = 0,$$

which implies that

$$\sin\left(\frac{\omega}{a} L\right) = 0.$$

In consequence, the argument of the sine function must have the values

$$\omega L/a = (n+1)\pi, \quad n = 0, 1, 2, 3, \dots, \quad (14.16)$$

where we have not considered the value of the argument to be equal to zero, because for that value $y(x, t) = 0$ for all values of x . Then, the frequency can only have certain values, namely,

$$\omega_n = \frac{\pi(n+1)}{L} \left(\frac{T}{\rho} \right)^{\frac{1}{2}} \quad n = 0, 1, 2, 3, \dots \quad (14.17)$$

Since for each value of ω_n there exists a solution of (14.12), there is an infinite number of solutions that, according to (14.15), can be written as

$$y_n(x, t) = B_n \sin\left(\frac{(n+1)\pi ax}{L}\right) [C_n \cos(\omega_n t) + D_n \sin(\omega_n t)]. \quad (14.18)$$

If we add the condition that, for $t = 0$, the string is at rest, ($\partial y / \partial t = 0$), and has an initial configuration $y(x, 0) = f(x)$, then equation (14.18) can be written as

$$y_n(x, t) = F_n \sin(k_n x) \cos(\omega_n t), \quad (14.19)$$

where $F_n = f(x) / \sin[\pi(n+1)x/L]$ and $k_n = (n+1)\pi/L$, where ω_n are the frequencies of vibration and k_n are the corresponding wave numbers. Equation (14.19) is analogous to (14.7) in the problem of standing waves with null values at $x = 0$ and $x = L$. Here also the introduction of a finite length of the string L leads to limiting the frequencies to a particular set. The general solution is given by the sum of solution (14.18) or (14.19) each corresponding to a particular frequency ω_n for all values of n . Therefore, the final solution of the problem of the vibrating string is a normal mode solution, that is, a sum of an infinite number of functions $y_n(x, t)$, each representing a harmonic motion of a different frequency ω_n , that is, a *mode of vibration*:

$$y(x, t) = \sum_{n=0}^{\infty} F_n \sin(k_n x) \cos(\omega_n t). \quad (14.20)$$

The mode corresponding to the lowest frequency ($n = 0$) is the fundamental mode,

$$\omega_0 = \frac{\pi}{L} \left(\frac{T}{\rho} \right)^{1/2}, \quad (14.21)$$

and its period is

$$T_0 = 2L \left(\frac{\rho}{T} \right)^{1/2} \quad (14.22)$$

The other values of n correspond to the *higher modes*, *harmonics*, or *overtones* whose frequencies are multiplied by $n+1$ and whose periods are divided by $n+1$.

Regarding the form that the string takes in its vibrations, it is easy to prove that, for the fundamental mode, there is no value of x in the interval $(0, L)$ for which $y(x, t)$ is zero; that is, there are no nodes of motion. For each higher mode there is a number of nodes equal to the

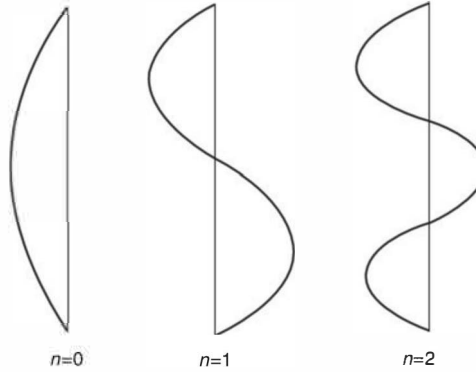


Fig. 14.3

The shape of a vibrating elastic string for the fundamental ($n = 0$) and first two higher modes ($n = 1, 2$).

order number of the mode (Fig. 14.3). The x coordinate of the position of the nodes for each mode of order n is given by

$$x_n^m = \frac{mL}{n+1}; \quad m = 1, 2, 3, \dots, n. \quad (14.23)$$

In conclusion, the problem of the vibration of an elastic string of finite length results in a solution with an infinite number of modes of vibration (14.19). The solutions $y_n(x, t)$ corresponding to each mode of vibration are the *eigenfunctions* of the system and the frequencies ω_n are the *eigenvalues* (eigenfrequencies). The complete solution of the problem is given by the sum of all modes (14.20).

14.3 Vibrations of an elastic rod

Another example of motion in a finite body is the vibration of an elastic rod of finite length. This problem introduces us, in a simple way, to the vibrations of an elastic medium of finite dimensions where we have longitudinal and transverse motion (corresponding to P and S wave motion). Let us consider a cylindrical rod of radius a and length L ($L \gg a$) and use cylindrical coordinates (r, z, ϕ) (Appendix 2) (Fig. 14.4).

14.3.1 Longitudinal vibrations

If we apply to the elastic rod a force in the direction of its axis (z), it will start to vibrate longitudinally along its axis and will continue vibrating after the force has ceased to act. Let us consider these vibrations after the force has been removed. If the material is isotropic, the relation between the stress and the derivatives of displacement (4.20), assuming that there are displacements only along the axis ($u_r = u_\phi = 0$ and $u_z(z, t)$), is

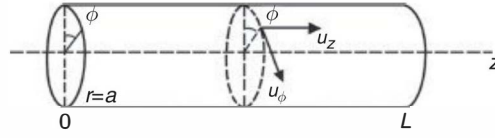


Fig. 14.4

A vibrating elastic rod: cylindrical coordinates and components of the displacements u_z and u_ϕ are shown.

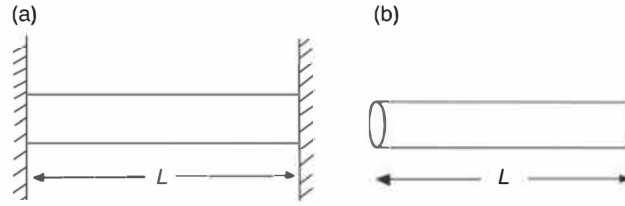


Fig. 14.5

A vibrating elastic rod: (a) with both ends fixed to a rigid medium; and (b) with both ends free.

$$\tau_{zz} = (\lambda + 2\mu) \frac{\partial u_z}{\partial z}. \quad (14.24)$$

The equation of motion in the absence of forces has the form of the wave equation (5.9),

$$\frac{\partial^2 u_z}{\partial z^2} = \frac{1}{\alpha^2} \frac{\partial^2 u_z}{\partial t^2}, \quad (14.25)$$

where $\alpha^2 = (\lambda + 2\mu)/\rho$ (4.72) is the velocity of longitudinal waves (P waves). The solution of this equation in the form of (14.3) may be written in two ways:

$$u_z = A \sin(kz) \cos(\omega t + \varepsilon), \quad (14.26)$$

$$u_z = A \cos(kz) \sin(\omega t + \varepsilon). \quad (14.27)$$

First, we will consider that both ends of the rod are fixed to a rigid material, in such a way that displacements at them are null ($u_z(0, t) = u_z(L, t) = 0$) (Fig. 14.5a). In this case, we use the solution (14.26), since it satisfies $u_z(0, t) = 0$. For the other end, the condition $u(L, t) = 0$ gives

$$A \sin(kL) = 0; \quad kL = (n + 1)\pi, \quad n = 0, 1, 2, 3, \dots \quad (14.28)$$

In consequence, as in the case of the vibrating string (14.18), the solution is given in terms of normal modes:

$$u_z''(z, t) = A_n \sin\left(\frac{(n + 1)\pi z}{L}\right) \cos(\omega_n t + \varepsilon_n); \quad n = 0, 1, 2, \dots, \quad (14.29)$$

where, just like in (14.17),

$$\omega_n = \frac{\pi(n + 1)\alpha}{L} \quad n = 0, 1, 2, 3, \dots \quad (14.30)$$

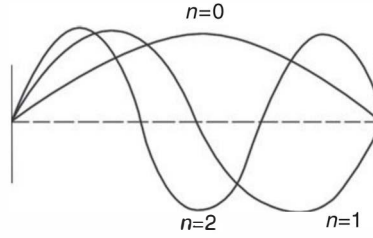


Fig. 14.6

Longitudinal vibrations of an elastic rod with both ends fixed, for the fundamental mode and the first two higher modes.

For each value of n , the solution (14.29) represents a mode of vibration. In this case, because of the type of motion, they are called *longitudinal modes*. The total motion is given by the sum of all modes. The lowest order mode, corresponding to $n = 0$, is the fundamental mode, whose frequency is $\omega_0 = \pi\alpha/L$; its wave length is $\lambda_0 = 2L$ and its period is $T_0 = 2L/\alpha$ (the time that it takes a longitudinal wave to travel along the rod in both directions). Frequencies of higher modes are multiples and periods of higher modes are fractions of those of the fundamental mode. For the fundamental mode there is no value of z along the rod where u_z is null. For each higher mode of order n , as in the case of the string, there exist n values of z , between 0 and L , where u_z is zero. These are given by the same equation (14.23) (Fig. 14.6).

Another possibility is that the two ends of the rod are free (Fig. 14.5b). The boundary conditions for free surfaces are that stresses through them (tractions) are null. In this case, we consider only normal components of the stress ($\tau_{zz}(0) = \tau_{zz}(L) = 0$). Now we take solution (14.27) and normal stress (if $e_{rr} = e_{\phi\phi} = 0$), according to (4.20) and (4.72), is given by

$$\tau_{zz} = -\rho\alpha^2 Ak \cos(kz) \sin(\omega t + \varepsilon). \quad (14.31)$$

The condition of null stress at $z = 0$ and $z = L$ leads to the same values of the wave numbers k_n and frequencies ω_n as in the previous case: (14.28) and (14.30).

According to (14.27), for all modes, the amplitude of displacements at each free end of the rod is always a maximum (for the fundamental mode, $n = 0$, $u_z(0) = A$, and $u_z(L) = -A$). The relation for the position of nodes along the rod's axis for different modes is now given (Fig. 14.7) by

$$z_n^m = \frac{(2m+1)L}{2(n+1)}; \quad m = 0, 1, 2, 3, \dots, n. \quad (14.32)$$

For the fundamental mode there exists a node at $z = L/2$. The number of nodes for each mode of order n is $n + 1$. Since we have not considered the applied force, in both cases the solutions correspond to free longitudinal vibrations of the elastic rod.

14.3.2 Torsional vibrations

Another way of generating vibrations in an elastic rod is to apply a torsional moment at a tangent to its surface and normal to its axis. The result is a torsion of the rod with

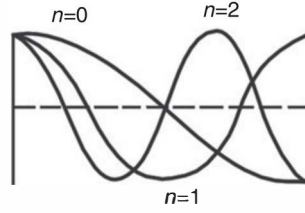


Fig. 14.7

Longitudinal vibrations of an elastic rod with both ends free, for the fundamental mode and the first two higher modes.

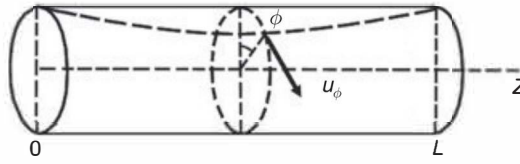


Fig. 14.8

Torsional vibrations of an elastic rod.

deformations in the direction of the angle ϕ , corresponding to displacements u_ϕ (Fig. 14.8). If u_ϕ varies only with z along the rod, the equation of motion is given by

$$\frac{\partial^2 u_\phi}{\partial z^2} = \frac{1}{\beta^2} \frac{\partial^2 u_\phi}{\partial t^2}, \quad (14.32)$$

where, just like in (4.73), $\beta^2 = \mu/\rho$ is the velocity of transverse waves (S waves). If the two ends of the rod are fixed to a rigid body so that its displacements are null ($u_\phi(0, t) = u_\phi(L, t) = 0$), then the solution of equation (14.32), after imposing the boundary conditions, as in the previous case, is given by

$$u_\phi^n(z, t) = A_n \sin\left(\frac{(n+1)\pi z}{L}\right) \cos(\omega_n t + \varepsilon_n); \quad n = 0, 1, 2, 3, \dots, \quad (14.33)$$

where

$$\omega_n = \frac{(n+1)\pi\beta}{L}; \quad n = 0, 1, 2, 3, \dots \quad (14.34)$$

For each value of n , there exists a mode of vibration with frequency ω_n . The value $n = 0$ corresponds to the fundamental mode and the others correspond to higher modes. Since the elastic deformation is of torsional character, these modes are called *torsional modes*.

In conclusion, just like the infinite medium, a finite isotropic elastic body has different responses to longitudinal and torsional stresses. Free vibrations of an isotropic elastic rod of finite length, with the simplifying assumptions that we have used, are formed by modes of two types, longitudinal and torsional, which are related to longitudinal (P) and transversal (S) waves. The lowest frequency corresponds to the fundamental mode and the frequencies for higher modes are multiples of it. The total motion is given by the sum of all modes.

14.4 The general problem. The Sturm–Liouville equation

The vibrations of an elastic rod of finite length is a particularly simple case of the general problem of the vibration of an elastic body of finite dimensions with arbitrary shape and boundary conditions. In general, problems of vibrations of elastic bodies do not have simple solutions such as those we have seen for the string and rod. The solutions depend on the geometry of the shape of the body and its dimensions and the boundary conditions on its external surface. In many cases, however, the problem can be reduced to equations of a particular type, known as the *Sturm–Liouville equation*, first proposed by Charles François Sturm and Joseph Liouville in 1837. This differential equation for a function $f(z)$ has the general form

$$\frac{d}{dz} \left(p(z) \frac{df(z)}{dz} \right) + [q(z) + ns(z)]f(z) = 0, \quad (14.35)$$

where $p(z)$, $q(z)$, and $s(z)$ are algebraic functions, commonly as quotients with finite numbers of roots of the numerator (zeros) and of the denominator (poles). The solutions of $f(z)$ for specific $p(z)$, $q(z)$, and $s(z)$ depend on the value of the parameter n and on the boundary conditions imposed on the problem. The most general properties of this problem are the following. For each value of n there is a solution $f_n(z)$, which is called an *eigenfunction*, and the general solution is the sum of these functions:

$$f(z) = \sum_n f_n(z). \quad (14.36)$$

There is a minimum value of n , but not a maximum value. As n increases, zeros of $f_n(z)$ correspond to values of z that are more similar to each other. When, for the same value of n , there are several different eigenfunctions $f_n(z)$, this is known as a *degenerate case*. The eigenfunctions $f_n(z)$ for different values of n form a complete set of orthogonal functions. If the functions are normalized, they are orthogonal and it follows that

$$\int_{-\infty}^{\infty} f_n(z) f_m(z) dz = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases}. \quad (14.37)$$

In the cases that we have considered for the elastic string and rod, equations (14.11), (14.24), and (14.32), applying the separation of variables as in (14.12) or assuming a harmonic dependence on time, results in equations like

$$c^2 \frac{d^2 f}{dz^2} + \omega^2 f = 0. \quad (14.38)$$

This is the simplest form of the Sturm–Liouville equation, in which $p(z) = c^2$, $q(z) = 0$, and $ns(z) = \omega^2$. The solutions f_n resulting from the application of boundary conditions are the eigenfunctions of the problem and the frequencies ω_n are its eigenvalues. As we have seen, the eigenvalues are real and the eigenfunctions, which are harmonic functions, are orthogonal and form a complete set. Another problem that can be solved with the help of

Sturm–Liouville equations are the modes of vibration of the vertical displacements of a thin membrane, held in a rectangular frame of finite dimensions. We will see now how these equations appear in the solutions of the free vibrations of a liquid and elastic sphere of finite radius.

14.5 Free oscillations of a homogeneous liquid sphere

The first approximation to the application of this approach to free oscillations or vibrations of the Earth is that of a homogeneous, non-rotating and non-gravitating, perfectly elastic sphere with the same radius, density, and elastic properties as the average values for the Earth. Further problems will introduce variations of density and elastic properties with the radius, more heterogeneous conditions, a lack of sphericity, and anelastic properties. As an introduction to the problem, let us first consider the simpler case of the free oscillations of a non-gravitating and non-rotating liquid sphere with constant density and bulk modulus. The components of displacements in spherical coordinates ([Appendix 2](#)) u_r , u_θ , and u_ϕ can be derived from a single scalar potential $\Phi(r, \theta, \phi, t)$ ([Section 5.9](#)) ([Fig. 14.9](#)):

$$u_r = \frac{\partial \Phi}{\partial r}, \quad (14.39)$$

$$u_\theta = \frac{1}{r} \frac{\partial \Phi}{\partial \theta}, \quad (14.40)$$

$$u_\phi = \frac{1}{r \sin \theta} \frac{\partial \Phi}{\partial \phi}. \quad (14.41)$$

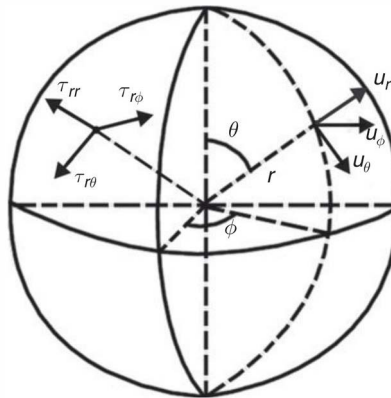


Fig. 14.9

Spherical coordinates and components of the displacements and stresses.

In a liquid medium the equation of motion in the absence of body forces (Section 5.1) leads to the potential Φ satisfying the wave equation

$$\nabla^2 \Phi = \frac{1}{\alpha^2} \frac{\partial^2 \Phi}{\partial t^2}, \quad (14.42)$$

where $\alpha^2 = K/\rho$ (K is the bulk modulus and ρ is the density, equation (4.73) with $\mu = 0$) is the velocity of longitudinal (acoustic) waves. If Φ has a harmonic time dependence,

$$\Phi(r, \theta, \phi, t) = \Phi(r, \theta, \phi) e^{-i\omega t}, \quad (14.43)$$

then equation (14.42) becomes the Helmholtz equation (5.10),

$$(\nabla^2 + k^2)\Phi = 0, \quad (14.44)$$

where $k = \omega/\alpha$ is the wave number. By expressing the Laplacian in spherical coordinates (A2.30) and applying the method of separation of variables, $\Phi(r, \theta, \phi) = R(r)N(\theta)L(\phi)$, we obtain three equations for R , N , and L , just like (5.113) to (5.115), where the constants of separation of variables are l and m (in Section 5.9, we used n and m). These three equations, upon introducing into the second equation the change of variable $y = \cos\theta$, may be written in the following form:

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + [r^2 k^2 - l(l+1)]R = 0, \quad (14.45)$$

$$\frac{d}{dy} \left((1-y^2) \frac{dN}{dy} \right) + \left(l(l+1) - \frac{m^2}{1-y^2} \right) N = 0, \quad (14.46)$$

$$\frac{d^2 L}{d\phi^2} + m^2 L = 0. \quad (14.47)$$

These three equations have the form of the Sturm–Liouville equation (14.35). The expressions for $p(z)$, $q(z)$, and $ns(z)$ are the following:

In equation (14.45): $p(r) = r^2$, $q(r) = r^2 k^2$, and $ns(r) = l(l+1)$.

In equation (14.46): $p(y) = 1 - y^2$, $q(y) = -m^2/(1 - y^2)$, and $ns(y) = l(l+1)$.

In equation (14.47): $p(\phi) = 1$, $q(\phi) = m^2$, and $ns(\phi) = 0$.

As we have seen in Section 5.9 in the problem of spherical waves, the solutions for $R(r)$ (14.45) are given by spherical Bessel functions $j_l(kr)$, those for $N(\theta)$ (14.46) by associate Legendre functions $P_l^m(\cos\theta)$, and those for $L(\phi)$ (14.47) by harmonic functions $e^{\pm im\phi}$ (Appendix 3). In consequence, we obtain for the potential Φ a solution as a sum of normal modes,

$$\Phi(r, \theta, \phi) = \sum_{l=0}^{\infty} j_l(kr) \sum_{m=-l}^l Y_l^m(\theta, \phi), \quad (14.48)$$

$$Y_l^m = (-1)^m \left(\frac{2l+1}{4\pi} \frac{(1-m)!}{(1+m)!} \right) P_l^m(\cos\theta) [C_l^m \cos(m\phi) + S_l^m \sin(m\phi)]. \quad (14.49)$$

We impose the boundary condition that, for each point of the free surface of the sphere, the normal component of the stress or pressure is null. According to (4.22) and (4.86), for $r = a$,

$$P = -K\nabla^2\Phi = 0. \quad (14.50)$$

By substitution into (14.44), putting k in terms of ω , the boundary condition becomes

$$\frac{K\omega^2}{a^2}\Phi(a) = 0. \quad (14.51)$$

Since in the solution (14.48), the only part that depends on r is given by spherical Bessel functions, condition (14.51) implies that

$$\frac{K\omega^2}{a^2}j_l\left(\frac{\omega a}{\alpha}\right) = 0. \quad (14.52)$$

This condition depends on the roots of $j_l(x)$ (Appendix 3). Using Rayleigh's formula to express spherical Bessel functions by sines and cosines (A3.14), for the first values of $l = 0, 1$, and 2 , these are

$$j_0(x) = \frac{\sin x}{x}, \quad (14.53)$$

$$j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}, \quad (14.54)$$

$$j_2(x) = \left(\frac{3}{x^3} - \frac{1}{x}\right) \sin x - \frac{3}{x^2} \cos x. \quad (14.55)$$

Each of these functions has an infinite number of roots, which, for the first values of l , are

$$l = 0, \quad x = 1.0\pi, \quad 2.0\pi, \quad 3.0\pi, \quad \dots$$

$$l = 1, \quad x = 1.4303\pi, \quad 2.4590\pi, \quad 3.4709\pi, \quad \dots$$

$$l = 2, \quad x = 1.8346\pi, \quad 2.8950\pi, \quad 3.9226\pi, \quad \dots$$

For each value of l , we use the subindex n for each of these roots. For example, according to equation (14.52), the roots for $l = 0$ correspond to

$$\frac{\omega a}{\alpha} = (n+1)\pi, \quad n = 0, 1, 2, 3, \dots \quad (14.56)$$

Thus, normal modes impose conditions on the possible values of the frequency. Since these values depend on both subindexes, l and n , possible frequencies are designated by ${}_n\omega_l$. For $l = 0$, according to (14.56), the frequencies are

$${}_n\omega_0 = \frac{\alpha\pi}{a}(n+1), \quad (14.57)$$

and the periods are

$${}_nT_0 = \frac{2a}{\alpha}(n+1). \quad (14.58)$$

Just as we saw for an elastic string and a rod, a liquid sphere vibrates at certain fixed frequencies that depend on its material properties (K and ρ or α) and its radius (a). There is an infinite number of modes of vibration, each with its own frequency that depends on two subindexes, n and l . For each value of l , $n = 0$ corresponds to the fundamental mode and $n \geq 1$ correspond to higher modes, harmonics, or overtones. The potential for $l = 0$, according to (14.48) and (14.49), is given by

$${}_n\Phi_0 = C_0 \frac{\alpha}{{}_n\omega_0 r} \sin\left(\frac{{}_n\omega_0 r}{\alpha}\right). \quad (14.59)$$

In this case, the subindex m has only the value zero, and the potential does not depend on θ and ϕ . Then, according to (14.39), (14.40), and (14.41), displacements have only radial components (u_r), and these modes are called *radial modes*. According to (14.57) and (14.58), the frequencies and periods corresponding to the first three modes are

$$\begin{array}{ll} {}_0\omega_0 = \frac{\pi\alpha}{a} & {}_0T_0 = \frac{2a}{\alpha} \\ {}_1\omega_0 = \frac{2\pi\alpha}{a} & {}_1T_0 = \frac{a}{\alpha} \\ {}_2\omega_0 = \frac{3\pi\alpha}{a} & {}_2T_0 = \frac{2a}{3\alpha} \end{array}$$

The period of the fundamental mode ($n = 0$) ${}_0T_0$ is the longest possible and corresponds to the time that a wave takes to travel with a velocity α along the diameter of the sphere. For a sphere of the size of the Earth ($a = 6370$ km) with velocity $\alpha = 9$ km s⁻¹, its period ${}_0T_0$ equals 23.6 min.

For $l = 1$, we have three values of m , -1 , 0 , and 1 , and there are also n roots of $j_1(\omega a/\alpha)$. Since $P_n^{-m} = (-1)^m P_n^m$, the potential Φ is given by

$${}_n\Phi_1 = j_1({}_n\omega_1 r/\alpha)[C_1^0 \cos\theta + (C_1^1 - C_1^{-1})\sin\theta \cos\phi + (S_1^1 + S_1^{-1})\sin\theta \sin\phi]. \quad (14.60)$$

According to the roots of $j_1({}_n\omega_1 a/\alpha)$, we have the following frequencies for the fundamental and two first higher modes:

$$\begin{aligned} {}_0\omega_1 &= 1.4303\pi\alpha/a, \\ {}_1\omega_1 &= 2.4590\pi\alpha/a, \\ {}_2\omega_1 &= 3.4709\pi\alpha/a. \end{aligned} \quad (14.61)$$

These frequencies are higher than those corresponding to modes for $l = 0$. For $l = 1$, the fundamental mode ($n = 0$) does not exist, since it implies a change in the position of the center of mass of the sphere, which is not allowed for free vibrations. In general, the higher the mode order the higher the corresponding frequencies. Since, for $l \geq 1$, the potentials (14.48) and (14.49) depend on the three coordinates r , θ , and ϕ , and the displacements

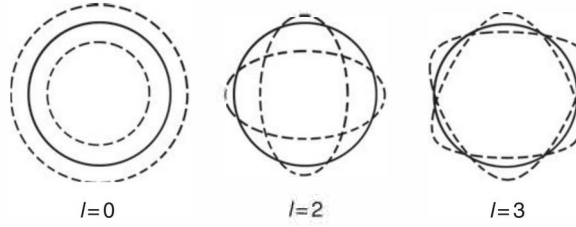


Fig. 14.10

Vibration modes in a liquid sphere for values of $l = 0, 2$, and 3 .

have three components, u_r , u_θ , and u_ϕ . For each value of l , m takes $2l + 1$ values and there are $2l + 1$ potential functions, all corresponding to the same frequency. For example, for $l = 2$, m takes the values $m = -2, -1, 0, 1$, and 2 , and there are five potentials for the same frequency ${}_n\omega_2$. This means that there are several eigenfunctions for one eigenvalue, which is known as a degeneracy problem. For general, nondegenerate cases, the frequencies depend on the three indexes n , l , and m (${}_n\omega_l^m$); modes depending on n are called *radial modes*, those depending on l are *angular modes*, and those depending on m are *azimuthal modes*.

If there is symmetry with respect to ϕ , that is for $m = 0$, then the displacements depend on θ , according to the Legendre polynomials $P_l(\cos \theta)$. For $l = 0$ ($P_0 = 1$), the displacements are purely radial (u_r). For $l = 2$, the body takes an ellipsoidal form, alternatively elongated and flattened at the poles ($\theta = 0$) with displacement components u_r and u_θ . For higher values of l , the forms are symmetric with respect to the equatorial plane ($\theta = \pi/2$) if l is even and asymmetric if l is odd (Fig. 14.10).

14.6 Free oscillations of an elastic sphere

A first approximation to the problem of free oscillations of the Earth is that of a homogeneous, perfectly elastic, non-rotating and non-gravitating sphere. This problem was first treated by Poisson in 1829 and in a more complete form by Lamb in 1882. In 1911, Love calculated the fundamental period of the vibrations of an elastic sphere with the dimensions of the Earth and found a value near to 1 h.

The problem is similar to that of the oscillations of a liquid sphere. Displacements in spherical coordinates are derived from a scalar potential Φ and a vector potential ψ , as we saw in Section 5.9. The vector potential ψ can be separated into two potentials that are products of the unit vector in the radial direction ($\mathbf{r}$), and a scalar potential S called the *spheroidal potential*, and another T called the *toroidal potential* (do not confuse with the period T):

$$\psi = \nabla \times S\mathbf{r} + T\mathbf{r}, \quad (14.62)$$

$$\mathbf{u} = \nabla\Phi + \nabla \times \nabla \times S\mathbf{r} + \nabla \times T\mathbf{r}. \quad (14.63)$$

In equation (14.63), the displacements are separated into three parts, the first term representing P wave motion, the second SV wave motion, and the third SH wave motion. In relation to surface waves, S represents Rayleigh wave motion and T represents Love

wave motion. When there is symmetry with respect to ϕ , the relations between these potentials and the scalar potentials Φ , ψ , and Λ , used in Section 5.9, are

$$\Phi = \Phi, \quad (14.64)$$

$$\psi = \frac{1}{r^2} \frac{\partial S}{\partial \theta}, \quad (14.65)$$

$$\Lambda = T/r. \quad (14.66)$$

The components of displacement in the direction of the coordinates r , θ , and ϕ in terms of the potentials Φ , S , and T are given by

$$u_r = \frac{\partial \Phi}{\partial r} + \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} \left(\frac{\sin \theta}{r} \frac{\partial S}{\partial \theta} \right) - \frac{1}{\sin \theta} \frac{\partial^2 S}{\partial \phi^2} \right], \quad (14.67)$$

$$u_\theta = \frac{1}{r} \frac{\partial \Phi}{\partial \theta} - \frac{1}{r} \frac{\partial}{\partial r} \left(\frac{\partial S}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial T}{\partial \phi}, \quad (14.68)$$

$$u_\phi = \frac{1}{r \sin \theta} \frac{\partial \Phi}{\partial \phi} + \frac{1}{r} \frac{\partial}{\partial r} \left(\frac{r}{\sin \theta} \frac{\partial S}{\partial \phi} \right) - \frac{1}{r} \frac{\partial T}{\partial \theta}. \quad (14.69)$$

Spheroidal motion depending on the potentials Φ and S is a combination of P and SV displacements with u_r , u_θ , and u_ϕ components, whereas toroidal motion depending on the potential T is of SH type with u_θ and u_ϕ components only.

The boundary conditions on the free surface ($r = a$) are that the components of the stress across it are null:

$$\tau_{rr} = \tau_{r\theta} = \tau_{r\phi} = 0. \quad (14.70)$$

For an isotropic medium, the components of the stress as a function of the displacement are given by

$$\tau_{rr} = (\lambda + 2\mu) \frac{\partial u_r}{\partial r} + \lambda \left[\frac{2u_r}{r} + \frac{1}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (u_\theta \sin \theta) + \frac{\partial u_\phi}{\partial \phi} \right) \right], \quad (14.71)$$

$$\tau_{r\theta} = 2\mu \left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} \right), \quad (14.72)$$

$$\tau_{r\phi} = 2\mu \left(\frac{1}{r \sin \theta} \frac{\partial u_r}{\partial \phi} + \frac{\partial u_\phi}{\partial r} - \frac{u_\phi}{r} \right), \quad (14.73)$$

where λ and μ are Lamé coefficients.

In the absence of body forces and without considering the effects of rotation and gravity, Navier's equation (4.75) is

$$\alpha^2 \nabla (\nabla \cdot \mathbf{u}) - \beta^2 \nabla \times \nabla \times \mathbf{u} = \ddot{\mathbf{u}}, \quad (14.74)$$

where α and β are the P and S waves velocities. The problem is solved by expressing the components of $\mathbf{u}$ in terms of the potentials Φ , S , and T , according to equations (14.67)–(14.69) and applying the surface boundary conditions (14.70). The complete problem is rather complex and we will give here only the most basic parts.

14.7 Toroidal modes

The part of the problem that corresponds to the toroidal potential $T(r, \theta, \phi, t)$ can be treated separately, as was done with SH waves (Chapter 5) and Love-wave propagation (Chapter 12). Substitution into Navier's equation (14.74) and assuming a harmonic dependence on time, we can write the Helmholtz equation for T and the problem is similar to that with a liquid sphere:

$$(\nabla^2 + k_\beta^2)T = 0, \quad (14.75)$$

where $k_\beta = \omega/\beta$ is the wave number of S wave motion. Just like in (14.48) and (14.49), the solution in spherical coordinates can be written in terms of spherical Bessel, associated Legendre, and harmonic functions:

$$T(r, \theta, \phi) = \sum_{l=0}^{\infty} j_l(k_\beta r) \sum_{m=-l}^l P_l^m(\cos \theta) [C_l^m \cos(m\phi) + S_l^m \sin(m\phi)]. \quad (14.76)$$

As mentioned before, the subindex l is the angular number and m is the azimuthal number. For each value of l there are $2l + 1$ values of m . The solution is given by an infinite sum of modes T_l^m that are called *toroidal modes*. Since the potential T represents transverse motion, these modes are similar to the torsional modes of vibration of an elastic rod (Section 14.3.2). From (14.68) and (14.69), the toroidal displacements are given by

$$u_\theta = -\frac{1}{\sin \theta} \frac{\partial T}{\partial \phi}, \quad (14.77)$$

$$u_\phi = -\frac{1}{r} \frac{\partial T}{\partial \theta}. \quad (14.78)$$

The boundary condition at the surface of the sphere ($r = a$) is $\tau_{r\phi} = 0$. From (14.73), since for toroidal motion $u_r = 0$, we obtain

$$\frac{\partial u_\phi}{\partial r} - \frac{u_\phi}{r} = 0, \quad r = a. \quad (14.79)$$

On putting u_ϕ as a function of T , according to (14.78), this condition gives

$$\frac{1}{r} \frac{\partial}{\partial r} \frac{\partial T}{\partial \theta} - \frac{1}{r^2} \frac{\partial T}{\partial \theta} = \frac{1}{r} \frac{\partial}{\partial \theta} \left(\frac{\partial T}{\partial r} - \frac{T}{r} \right) = 0, \quad r = a. \quad (14.80)$$

According to (14.76), T is a product of three parts, each depending on only one of the coordinates, $T(r, \theta, \phi) = T_r(r)T_\theta(\theta)T_\phi(\phi)$. Equation (14.80) imposes only a condition on $T_r(r)$ in the form

$$r \frac{\partial T_r}{\partial r} - T_r = 0, \quad r = a. \quad (14.81)$$

Since, according to (14.76), $T_r = j_l(k_\beta r)$, on putting $x = k_\beta r$, condition (14.81) gives

$$x j'_l(x) = j_l(x), \quad x = k_\beta a. \quad (14.82)$$

Using the relation $j'_l = [1/(l-1)]j_{l+1}(x)$, this condition may be written as

$$x j_{l+1}(x) = (l-1)j_l(x), \quad x = k_\beta a. \quad (14.83)$$

This equation has an infinite number of solutions for which we use the subindex n . For each value of l , the first value of n , that is, $n = 0$, corresponds to the fundamental mode and the rest to higher modes, harmonics, or overtones.

For $l = 0$, $P_\bullet(\cos \theta) = 1$ and $T_\bullet = j_\bullet(k_\beta r)$. Then, $\partial T / \partial \theta = \partial T / \partial \phi = 0$ and $u_\theta = u_\phi = 0$; that is, there are no toroidal vibrations of order zero.

For $l = 1$ ($P^0_1 = \cos \theta$ and $P^1_1 = \sin \theta$), condition (14.83) corresponds to the roots of $j_2(x)$, which, according to (14.55), gives

$$\tan x = \frac{3x}{3 - x^2}, \quad x = k_\beta a. \quad (14.84)$$

The first three roots, corresponding to $n = 0, 1$, and 2 , are $x = 1.8346\pi$, 2.8950π , and 3.9226π . Since $k_\beta = \omega/\beta$, the eigenfrequencies for $l = 1$ and $n = 0, 1$, and 2 are

$${}_0\omega_1 = \frac{1.8346\pi\beta}{a}; \quad {}_1\omega_1 = \frac{2.8950\pi\beta}{a}; \quad {}_2\omega_1 = \frac{3.9226\pi\beta}{a}.$$

For toroidal modes, a solution for $n = 0$ is not physically possible for free vibrations (in the absence of external torques) since this corresponds to a rigid oscillation of the whole sphere. For $n = 1$ and 2 , and $l = 1$, if we use approximate values of a and β for the Earth ($a = 6370$ km and $\beta = 6$ km s⁻¹), then the corresponding periods t (we are using T for the toroidal functions) in minutes are ${}_1t_1 = 12.22$ and ${}_2t_1 = 9.02$.

The solutions for the potential function $T(r, \theta, \phi)$ (in general ${}_nT^m_l$) for $l = 1$ and $m = -1, 0, 1$ is, according to (14.76),

$${}_nT_1 = j_1({}_n\omega_1 r/a) \{C^0_1 \cos \theta + \sin \theta [(C^1_1 - C^{-1}_1) \cos \phi + (S^1_1 + S^{-1}_1) \sin \phi]\}. \quad (14.85)$$

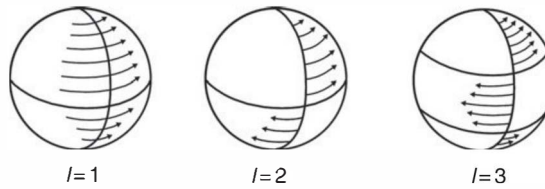
For each value of n , this can be considered as a sum of three eigenfunctions, ${}_nT^0_1$, ${}_nT^1_1$, and ${}_nT^{-1}_1$, that correspond to the same eigenvalue, ${}_n\omega_1$. Therefore, the problem is a degenerate one with several eigenfunctions for the same eigenvalue (Section 14.4).

For $l = 2$, condition (14.83), by substitution of (14.55), gives

$$\tan x = \frac{(12 - x^2)x}{12 - 5x^2}. \quad (14.86)$$

Table 14.1 Periods (in minutes) of toroidal modes.

l/n	0	1	2	3
1	—	13.45	7.63	5.22
2	43.82	12.60	7.49	5.17
3	28.35	11.56	7.28	5.08

**Fig. 14.11** Toroidal modes for values of $l = 1, 2$, and 3 .

The first three roots of this equation, corresponding to $n = 0, 1$, and 2 , are $x = 0.796\pi, 2.271\pi$, and 3.346π . The corresponding periods, for the given values of α and β , are ${}_0t_2 = 44.46$ min, ${}_1t_2 = 15.58$ min, and ${}_2t_2 = 10.58$ min. In this case, we have a fundamental mode (${}_0T_2$) which has the longest period (lowest frequency) of all toroidal modes. The period of mode ${}_2T_2$ (44.46 min) is greater than the time (35.39 min) which an SH wave takes to travel along the diameter of the sphere.

For toroidal modes, the index l refers to the order of associated Legendre functions and in consequence to the distribution of displacements on spherical surfaces depending on the angle θ . For $l = 1$, the fundamental mode implies that whole-body oscillations around the origin of the θ axis, which are not possible as free vibrations, occur as it will imply a change in total angular momentum. They are possible for higher modes since internal parts oscillate in different senses and there is no change in total angular momentum. For $l = 2$, the two hemispheres oscillate in opposite senses. For higher values ($l \geq 3$), there are as many zones oscillating in opposite senses as the order number of the mode (Fig. 14.11). The index m refers to the dependence of displacements on the azimuthal angle ϕ and takes values from -1 to 1 . The index n refers to the number of roots in the r dependence of displacements for each configuration of l and m (Fig. 14.12). Since, for a homogeneous sphere, the eigenfrequencies do not depend on the index m , displacements that are different according to their value of m correspond to the same frequencies; that is, the problem is a degenerate one. The same happens if elastic properties depend on the radius only, but not for the complete heterogeneous problem. For each value of l and m , $n = 0$ corresponds to the fundamental mode and $n \geq 1$ correspond to the higher modes or overtones.

The eigenperiods in minutes for toroidal modes of low order corresponding to a model of the Earth with radial symmetry are given in Table 14.1 (Dziewonski and Gilbert, 1972). Periods decrease with increasing mode-order number. The longest period, 43.82 min, corresponds to the fundamental mode of the second-order mode ${}_0T_2$. The values of the

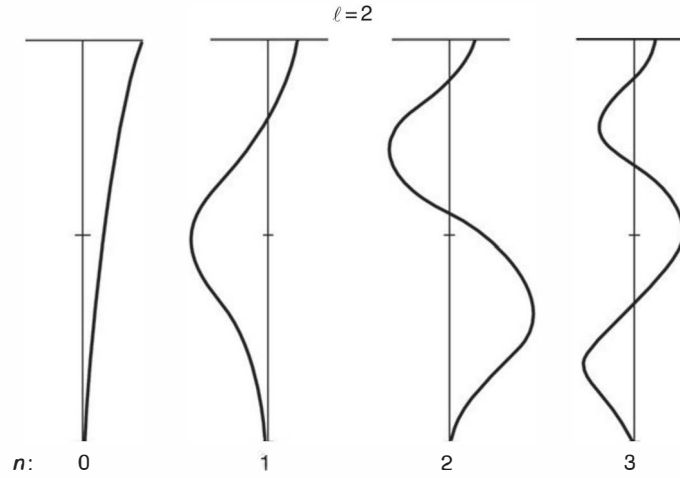


Fig. 14.12 Distribution of amplitude with radius for the fundamental ($n = 0$) and higher-order harmonics ($n = 1, 2$, and 3) for the toroidal mode of order $l = 2$.

periods in Table 14.1 are similar to those found for a homogeneous sphere with the assumed value of S wave velocity (6 km s^{-1}), for example, for mode ${}_1T_1$ the periods are 13.45 and 12.22 min and for mode ${}_2T_1$ they are 7.63 and 9.02 min. The displacements of the toroidal modes (u_θ and u_ϕ) are derived from the potential T according to (14.77) and (14.78). The total toroidal motion for a homogeneous elastic sphere is given by the sum of the different modes. Using the appropriate models, in this way we can also generate the motion corresponding to Love and SH waves. Motion including short periods requires the sum of a large number of modes.

14.8 Spheroidal modes

The problem of spheroidal modes is more complicated since it implies solutions for two potentials, Φ and S , and the displacements (u_r , u_θ , u_ϕ) correspond to P–SV motion. The relation between potentials Φ and S and displacements (u_r , u_θ , u_ϕ) are given by (14.67)–(14.69), putting $T = 0$. Let us consider briefly a homogeneous sphere of velocities α and β , and density ρ . The boundary conditions at the free surface are that the corresponding components of the stresses are null:

$$\tau_{rr} = \tau_{r\theta} = 0, \quad r = a. \quad (14.87)$$

For harmonic dependence on time, the equation of motion leads to Helmholtz equations, which in spherical coordinates have solutions for potentials Φ and S of the same type as those found for T (14.76). In a similar form as for toroidal modes ${}_nT_l$ different spheroidal modes are represented by ${}_nS_l$.

Table 14.2 Periods (in minutes) of spheroidal modes.

l/n	0	1	2	3
0	20.45	10.25	6.62	5.10
1	—	41.30	17.66	11.79
2	53.80	24.5	15.08	8.48
3	35.53	17.72	13.43	8.14

As in the case for the vibrations of a liquid sphere (14.59), for the lowest order $l = 0$, in the spheroidal modes the potentials Φ and S are functions of r only. Then, according to (14.67)–(14.69), there are only displacements u_r , and $u_\theta = u_\phi = 0$. These modes are, then, called *radial modes*. The boundary condition $\tau_{rr} = 0$ results in the equation (Ben Menahem and Singh, 1981)

$$\cot x = \frac{1}{x} - \frac{1}{4} \left(\frac{\alpha}{\beta} \right)^2 x, \quad (14.88)$$

where $x = k_a a$. For $\alpha = \sqrt{3}\beta$ ($\lambda = \mu$), the first roots are $x = 0.816\pi$, 1.929π , 2.936π , and 3.966π , corresponding to harmonics of orders $n = 0, 1, 2$, and 3 . If, for an approximation to the Earth, we substitute $a = 6370$ km and $\alpha = 9$ km s⁻¹, the corresponding periods in minutes are 28.91, 12.23, 8.03, and 5.95. The fundamental mode ${}_0S_0$ corresponds to an expansion and contraction of the sphere without its form changing, as we saw for the case of a liquid sphere (Fig. 14.10). For higher values of n , there are as many nodal surfaces inside the sphere as the order number for which the motion is null and changes sign.

For $l = 1$, the fundamental mode ${}_0S_1$ does not exist, since it corresponds to a displacement of the center of mass, which is not possible in free vibrations, as we saw for the liquid sphere (Lapwood and Usami, 1981). We have values for the higher modes ${}_1S_1$, ${}_2S_1$, ${}_3S_1$, etc.

For $l = 2$, displacements are symmetric with respect to the plane normal to the origin of the axis θ . The surface takes an ellipsoidal form that is alternately flattened and elongated at the poles (Fig. 14.10). The fundamental mode ${}_0S_2$ has the lowest frequency of all modes with a period of 53.80 min (Table 14.2) and no nodes of displacement in the interior of the sphere. For higher modes ($n \geq 1$), just like for toroidal modes, there are in the interior of the sphere as many nodal surfaces as the order of the mode.

Eigenperiods in minutes for spheroidal modes of lowest order for a model of the Earth with radial symmetry are given in Table 14.2 (Dziewonski and Gilbert, 1972). As already mentioned the longest period, 53.8 min, corresponds to the mode ${}_0S_2$ and is greater than 43.8 min, the period of the corresponding toroidal mode ${}_0T_2$. Thus, this is the longest period for free oscillations of the Earth. For the same order, periods of spheroidal modes are longer than those of toroidal modes. Total spheroidal motion for an homogeneous elastic sphere is given by the sum of all modes. Using the appropriate models Rayleigh wave and P–SV motion can be expressed as a sum of spheroidal modes. Motion at shorter periods requires sums of a large number of modes.

A way to provide a deeper understanding of the relation between the normal modes of free oscillations and physical properties of the medium is given by the possibility of linking their angular order l to the traveling surface and body waves. The essential quantitative connection between modes and traveling waves is made by equating the horizontal wave length (or wave number) of the mode with the corresponding horizontal wave length (or wave number) of a traveling wave. Accordingly, the wave number k of a spherical harmonic of degree l can be approximated as $k = \frac{(l + \frac{1}{2})}{a}$, where a is the sphere's radius. Thus, for a given mode an apparent phase velocity $c(\omega) = \omega/k$ can be identified at the Earth's surface as:

$$c(\omega) = \frac{\omega a}{(l + \frac{1}{2})}. \quad (14.89)$$

In the same manner, for the case of body waves, we can express the horizontal wave number in terms of frequency and ray parameters as $k = \frac{\omega p}{a}$. A mode of angular order l and frequency ω can therefore be associated with rays having $p = \frac{(l + \frac{1}{2})}{\omega}$. And, since rays exist only for ranges of depth for which $\frac{r}{\alpha(r)} \geq p$ or $\frac{r}{\beta(r)} \geq p$, these inequalities are then satisfied only at certain depths. Moreover, by making use of the so-called dispersion diagrams (i.e. theoretically computed frequencies, ${}_n\omega_l$ versus angular order number, l) it becomes possible to compare information about specific parts of the Earth's interior from body waves with those from normal modes that have specific n , l values.

The structure of the Earth's interior can also be obtained from observations of the modes of free oscillations by the use of *differential kernels*, $K(r)$, which represent the derivatives of the eigenfrequency of a mode with respect to structural changes at a value of the radius. For instance, differential kernels can be defined corresponding to perturbations in bulk and shear moduli at fixed density, and these are related to the distribution of elastic shear and compressional energy with radius. Assuming relative perturbations in velocities α and β , and density ρ , we can then apply the *variational method*, where perturbations in phase velocity (due to small changes in the structure) are replaced by perturbations in eigenfrequency:

$$\delta\omega_k = \int_0^a \left(K_\rho(r) \frac{\delta\rho(r)}{\rho(r)} + K_p(r) \frac{\delta\alpha(r)}{\alpha(r)} + K_s(r) \frac{\delta\beta(r)}{\beta(r)} \right) dr. \quad (14.90)$$

Therefore, differential kernel analysis tells us how the frequency of a given mode changes along with specific variations in velocity or density at a certain depth, offering the possibility of studying the structure of the Earth's interior.

14.9 Effects on free oscillations

We have considered only the properties of the free oscillations of an isotropic homogeneous elastic, non-rotating and non-gravitating sphere. This is only a first approximation to the

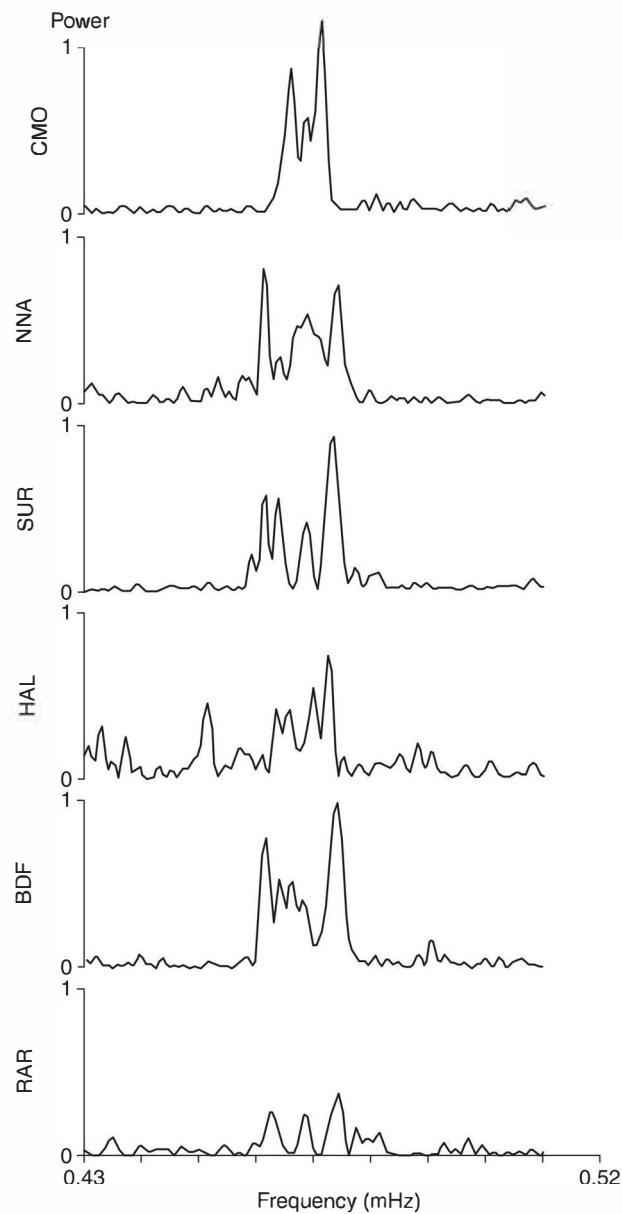


Fig. 14.13 Splitting of the spectral peak of mode ${}_0S_3$ from observations at six stations of the IDA network of the Indonesia earthquake of 19 August 1977 (Buland *et al.*, 1979) (with permission from Macmillan Magazine Ltd).

situation in the Earth. If the elastic properties vary continuously with radius the solutions are similar to those we have found for the homogeneous case. In the Earth there are circumstances that deviate from this simple model. First of all we have the effect of gravity, which we have not considered, and that affects mainly spheroidal modes. Secondly, the

presence of a liquid core results in the existence of toroidal modes in the mantle only. This influence is greater for low-order modes. For example, spheroidal modes have periods significantly different from those in models without a liquid core. These two effects do not change the spherical symmetry of the problem.

Three additional effects derive from the ellipticity of the shape of the Earth, its lateral heterogeneity (the dependence of θ and ϕ on the elastic properties), and its rotation. These three factors separate the problem from that of spherical symmetry. The problem is no longer degenerate and the values of the eigenfrequencies are modified. For each value of l and n , there are now several frequencies, depending on the index m with a maximum of $2l + 1$. These frequencies are called *multiplets* (Fig. 14.13).

The most important of these effects is due to the Earth's rotation. Centrifugal and Coriolis forces have axial symmetry, so the problem no longer has spherical symmetry. This effect produces splitting of the eigenfrequencies (Dahlen, 1980). If the angle ϕ is measured on a plane normal to the axis of the Earth's rotation, a perturbation, due principally to the Coriolis force, displaces the eigenfrequencies associated with modes of angular order l by a quantity related to the index m , $\delta\omega = mb$, where b is a function of the index l and the quotient Ω/ω_l (Ω is the Earth's frequency of angular rotation and ω_l are the unperturbed eigenfrequencies). Since this quotient is small for high eigenfrequencies, this effect is observed only for low frequencies.

Two other effects are due to the Earth's ellipticity and deviation of elastic properties and density from radial symmetry. These effects are small and also produce splitting of eigenfrequencies. If the lack of radial symmetry in the Earth's composition were very large, then, besides splitting of eigenfrequencies, we could have coupling of toroidal and spheroidal modes. Although their effects are small, heterogeneities in the Earth's composition can be detected in the analysis of free oscillations of high order.

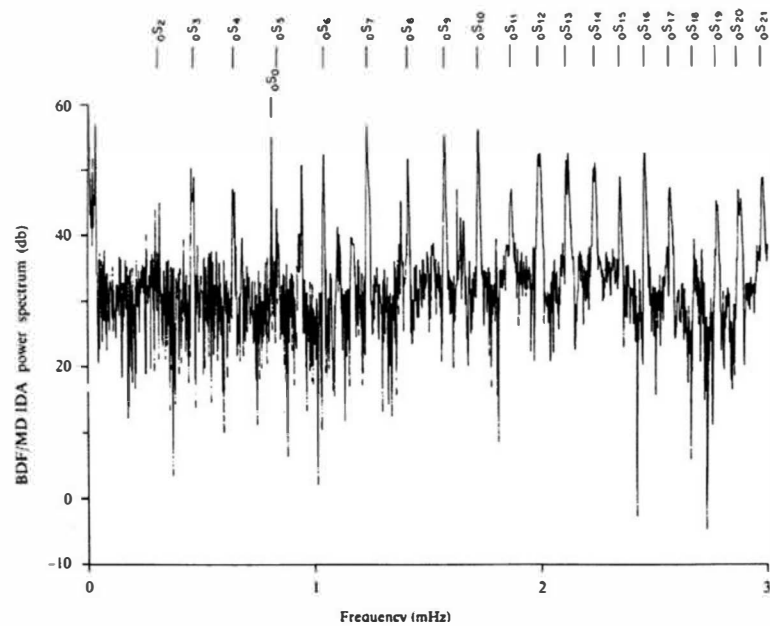
14.10 Observations

Free oscillations of the Earth can be observed in the analysis of seismograms of large earthquakes ($M > 7$). In large shocks, enough energy is released to generate low-frequency free oscillations that can be observed by means of broad band or long-period seismographs and gravity meters. However, the eigenfrequencies of the various modes are not observed directly on seismograms, but must be obtained from the peaks of their power spectra (Appendix 4). The eigenfrequencies of the various modes are found from observations taken from spectral analysis of long-period records using long time windows in order to obtain sufficient precision for peaks of the spectra (Fig. 14.14). To increase the signal-to-noise ratio, methods involving the stacking of several records for the same earthquake are used to correct for phase differences.

The first observations of free oscillations of the Earth were obtained from long-period records of the two large ($M > 8$) earthquakes in Kamchatka in 1952 and

Table 14.3 Observed periods (in minutes) of low-order modes.

1	${}_0T_1$	${}_0S_1$
2	44.011	53.883
3	28.463	35.559
4	21.739	25.786
5	17.938	19.821
6	15.422	16.065

**Fig. 14.14** Spectral peaks of spheroidal modes observed at the BDF station of the IDA network (Buland *et al.*, 1979) (with permission from Macmillan Magazine Ltd).

Mongolia in 1957, but the earthquakes that provided better data were those in Chile in 1960, the Kurile Islands in 1964, and Alaska in 1964 (Alsop *et al.*, 1961; Benioff *et al.*, 1961); and the recent ones in Sumatra in 2004 and Japan in 2011 (Zürn *et al.*, 2015). Splitting of spectral peaks due to the rotation of the Earth, especially for low-order modes, was observed even in the earliest studies. Observed values of periods in minutes for low-order toroidal and spheroidal fundamental modes are given in Table 14.3 (Derr, 1969).

Comparison of observed values of eigenperiods of free oscillations and those calculated from theoretical models is used to obtain global models of the Earth's interior. If we compare corresponding values of Table 14.3 with those of Tables 14.1 and 14.2, we find

that the model with radial symmetry is not a bad approximation for very-low-order modes. Since the greatest energy for each mode corresponds to a particular depth, they provide information about the structure at such a depth. Thus, mode data may contribute to the determination of the structures in the Earth at depths for which the structure is not well known from body-wave data. For example, mode analysis gives a value of S wave velocity of 3.5 km s^{-1} for the rigidity of the inner core. The splitting of spectral peaks due to lateral heterogeneities is also used in mode studies of the Earth's interior. Modern models of the Earth's interior integrate data from body waves, surface waves, and free oscillations covering a wide range of frequencies.

14.11 Summary

The study of motion with long wave lengths must take into consideration the finite dimensions of the Earth. This leads to the approach known as normal mode theory. First we have considered one-dimensional standing waves, introducing a finite length that had led to the solution as the sum of modes of vibrations of discrete frequencies. The vibrations of an elastic string of finite length show that the solution is also given by a sum of discrete frequencies (14.17), the lowest ($n = 0$) corresponding to the fundamental mode and the rest to higher-order modes. The problem of the vibrations of an elastic rod has introduced the existence in an elastic body of longitudinal and torsional vibrations depending on P and S wave velocities. Thus we have in an elastic medium two types of normal mode, longitudinal and torsional.

As a first approximation, the Earth can be considered as a homogeneous non-rotational and non-gravitating elastic sphere of finite radius. Therefore the problem has to be solved in spherical coordinates. As an introduction we have presented the solution for the free oscillations of a homogeneous liquid sphere. The solution for the potential is expressed in terms of spherical Bessel functions and associate Legendre functions. There is an infinite number of normal modes each with their own frequency that now depends on two subindexes (n and l). The lowest mode ($n = l = 0$) has a period corresponding to the time a P wave takes to travel the diameter of the sphere. Free oscillations in an elastic sphere are separated into toroidal and spheroidal modes. Toroidal modes ${}_nT_l$ have SH motion and spheroidal ${}_nS_l$ P–SV motion. The longest period 53.8 s corresponds to the fundamental spheroidal mode ${}_0S_2$.

In the Earth we must consider conditions which deviate from the simple model of the homogeneous elastic sphere, such as its rotation, variations of elastic properties and density in its interior, the presence of a liquid outer core, gravity, and its ellipticity of shape. For example, rotation and lateral heterogeneities lead to splitting of the eigenfrequencies. Free oscillations have been observed from long-period motion recorded in large earthquakes. Comparison of observed values of eigenperiods of free oscillations and those calculated from theoretical models, together with observations of body and surface waves, are used to obtain global models of the structure and composition of the Earth's interior.

14.12 Problems

- P14.1 Consider an elastic rod with an end fixed to a rigid wall and the other end free. Write the solutions for torsional vibrations and the values of the eigenfrequencies.
- P14.2 For a perfect elastic sphere of radius 6000 km, density 5.5 g/cm^3 , and rigidity $15 \times 10^4 \text{ MPa}$, calculate the value of the period of the torsional mode ${}_0T_2$ and compare with the observed values in the Earth.
- P14.3 For a perfect elastic sphere of radius 6000 km, P wave velocity 10 km/s, and Poisson ratio 0.25, calculate the lowest frequencies (largest periods) for spheroidal and torsional modes. Compare the periods with the times for P and S waves to travel inside the Earth across the diameter.

In all previous chapters we have considered the mechanical behavior of perfectly elastic bodies to approximate the motion produced by earthquakes in the Earth. Many problems in seismology can be solved assuming for the Earth an approximation of perfect elasticity and isotropy. We know, however, that the Earth cannot behave in this way and in some cases is important to take in account the anelasticity and anisotropy of its material.

Perfectly elastic bodies do not exist since according to the second law of thermodynamics, in all physical deformable media, there is dissipation of energy in the form of heat through mechanisms known as *internal friction* (Section 4.4). As we have seen, in a perfectly elastic medium, the amplitudes of waves only decrease with distance due to geometric spreading. The lack of perfect elasticity adds a further decrease in wave amplitude due to *anelastic attenuation*. From the very first studies (Jeffreys, 1957; Lomnitz, 1957), anelastic attenuation of seismic waves in the Earth has been an important subject of seismology (Jackson and Anderson, 1970; Minster, 1980; Romanowicz and Mitchell, 2009).

In addition to perfect elasticity, we have also assumed isotropy for the material of the Earth. Now we will consider how this material is not only perfectly elastic, but also it is not perfectly isotropic. Consideration of the lack of isotropy or *anisotropy* in the Earth is an important subject in seismology and it is related to geodynamic processes that influence its material properties. Although deviations from the conditions of isotropy are small, their effect on the propagation of seismic waves can be observed and provide important information about the Earth's structure and dynamics (Crampin, 1977; Babuska and Cara, 1991; Maupin and Park, 2009).

15.1 Anelasticity and damping

Loss of energy in the Earth due to internal friction is responsible for the attenuation and damping of the amplitudes of seismic waves with distance and time. Concrete mechanisms of internal friction are complex and depend on the atomic and molecular structure of crystals in minerals, the presence of small cracks and fractures, and the inclusion of liquids in rocks. Since these mechanisms depend on the nature of the materials through which waves propagate, their effect is called *intrinsic attenuation*.

15.1.1 Anelasticity

According to Hooke's law, the deformation (strain) of a perfectly elastic body is proportional to the applied stress (4.16). Once stresses are removed, the body instantaneously recovers its initial form. In one dimension, an elastic body may be represented by a spring and the strain–stress relation is given by (Fig. 15.1 a)

$$\sigma = \mu e, \quad (15.1)$$

where σ represents the stress, e is the elongation or strain, and μ is the coefficient of elasticity of the spring.

A different type of mechanical behavior corresponds to a *viscous body* for which, according to Stokes' law (presented by George Gabriel Stokes in 1851), the applied stress is proportional to the time derivative of the strain. In one dimension, a viscous body can be represented by a dashpot with viscosity coefficient η , and the relation between the stress and the strain (Fig. 15.1 b) is

$$\sigma = \eta \frac{de}{dt} = \eta \dot{e}. \quad (15.2)$$

Imperfectly elastic bodies can be considered as having properties intermediate between those of elastic and viscous bodies, and are called *viscoelastic bodies*. Their behavior can be understood in terms of one-dimensional mechanical models consisting of combinations of springs and dashpots. A *Maxwellian body* (named after James Clerk Maxwell) consists of an elastic (spring) and a viscous (dashpot) element in series (Fig. 15.2a). If a stress σ acts on both elements, the deformations in each of them are different and given by

$$\sigma = \mu e_1, \quad (15.3)$$

$$\sigma = \eta \dot{e}_2. \quad (15.4)$$

The total deformation of the system is $e = e_1 + e_2$, and its rate is given by

$$\dot{e} = \frac{\dot{\sigma}}{\mu} + \frac{\sigma}{\eta}. \quad (15.5)$$

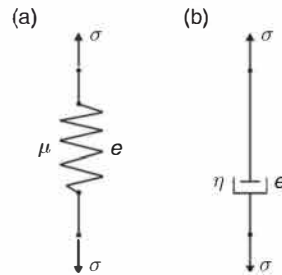


Fig. 15.1

Mechanical models of elastic and viscous bodies: (a) elastic (a spring) and (b) viscous (a dashpot).

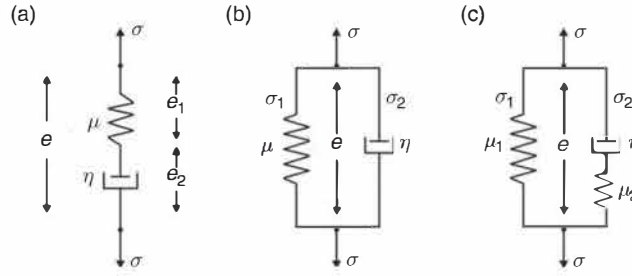


Fig. 15.2

Combinations of springs and dashpots to represent viscoelastic bodies: (a) in series – a Maxwell body; (b) in parallel – a Kelvin–Voigt body; (c) a standard linear solid.

A *Kelvin–Voigt body* (named after Lord Kelvin and Woldemar Voigt) is represented by an elastic (spring) and a viscous (dashpot) element in parallel (Fig. 15.2b). The deformation in the spring and dashpot is the same e , but the stresses acting on each element are different. The total stress is the sum $\sigma = \sigma_1 + \sigma_2$, namely

$$\sigma = \mu e + \eta \dot{e}. \quad (15.6)$$

In both models, when a stress is applied, the spring is deformed immediately but the dashpot takes some time to respond. Their combined action results in the system presenting a time delay of the elastic deformation with respect to the applied stress, an *inelastic deformation*, and a *relaxation time* for the system to recover its initial state when the stress is removed. The anelastic behavior of the Earth's materials does not agree with either of the two simple models, and more complex systems with more than two elements have been proposed. One of these is the so-called *standard linear solid* or *Zener model* (after Clarence M. Zener), consisting of a combination of three elements, two springs and a dashpot (a Kelvin–Voigt element plus a spring, Fig. 15.2c), which results in a more realistic representation of viscoelastic materials. A more complex model is a combination of several elements of the standard linear solid in series or parallel, called a *general standard linear solid*.

15.1.2 Harmonic excitation of a Maxwellian body

Certain properties of the response of imperfect elastic bodies may be understood by studying the excitation of a Maxwellian body by a harmonic stress:

$$\sigma = \sigma_0 \sin(\omega t). \quad (15.7)$$

On substituting into equation (15.5) and integrating, we obtain for the deformation at a time t ($0 < t < \pi/\omega$),

$$e(t) = \frac{\omega \sigma_0}{\mu} \int_0^t \cos(\omega \tau) d\tau + \frac{\sigma_0}{\eta} \int_0^t \sin(\omega \tau) d\tau. \quad (15.8)$$

Solving the integrals gives

$$e(t) = \frac{\sigma_0}{\mu} \left\{ \left[1 + \left(\frac{\mu}{\eta\omega} \right)^2 \right]^{1/2} \sin(\omega t + \phi) + \frac{\mu}{\eta\omega} \right\}, \quad (15.9)$$

where $\phi = \tan^{-1}[\mu/(\eta\omega)]$. A perfect elastic response (for one spring only) is

$$e(t) = \frac{\sigma_0}{\mu} \sin(\omega t). \quad (15.10)$$

On comparing (15.9) and (15.10) we see the difference between the response of Maxwellian and pure elastic bodies. This difference can be expressed by a parameter Q' , defined as

$$Q' = \frac{\mu}{\eta\omega}. \quad (15.11)$$

Equation (15.9) becomes

$$e(t) = \frac{\sigma_0}{\mu} [(1 + Q'^2)^{1/2} \sin(\omega t + \phi) + Q'], \quad (15.12)$$

where $\phi = \tan^{-1} Q'$. Thus the parameter Q' represents how much the response of a Maxwellian body differs from that of a perfectly elastic one. If it is zero, the problem reduces to one of perfect elasticity (15.10). The presence of the element of viscosity (value of Q') modifies the amplitude of the deformation and introduces a phase shift between the stress and the strain (Fig. 15.3).

15.1.3 Damped harmonic motion. The Q coefficient

The effect of anelastic deformation on wave propagation can be understood by considering *damped harmonic motion*. Let us consider a mass m , suspended by a spring of elasticity coefficient μ , and a dashpot of viscosity η mounted in parallel, in a configuration similar to that of a Kelvin–Voigt body (Fig. 15.4). The equation of motion

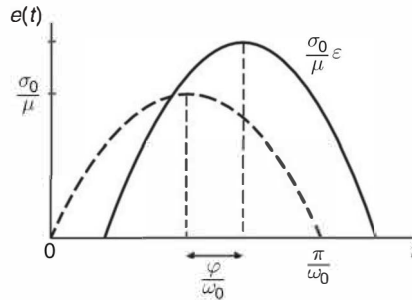


Fig. 15.3

The response of a Maxwell body to a harmonic excitation of frequency ω_0 . The dashed line is for an elastic body.

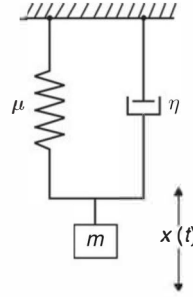


Fig. 15.4 A mechanical model of a system with damped harmonic motion.

of the mass is given by (we have already seen this equation in [Section 3.2](#), about the theory of the seismometer)

$$\ddot{x} + \frac{\eta}{m} \dot{x} + \frac{\mu}{m} x = 0. \quad (15.13)$$

In a system without damping (one with no dashpot), the motion is harmonic with a frequency $\omega_0 = (\mu/m)^{1/2}$. The damping effect of the dashpot can be represented by the coefficient Q , which is now defined by

$$\frac{1}{Q} = \frac{\eta}{m\omega_0}. \quad (15.14)$$

Its relation with Q' of the Maxwellian body is

$$Q = \omega \sqrt{\frac{m}{\mu}} Q'.$$

On replacing the values of ω_0 and Q , we rewrite equation (15.13) as

$$\ddot{x} + \frac{\omega_0}{Q} \dot{x} + \omega_0^2 x = 0. \quad (15.15)$$

The solution of this equation is

$$x(t) = A \exp\left(-\frac{\omega_0}{2Q} t\right) \sin(\omega t + \phi), \quad (15.16)$$

$$\omega = \left(1 - \frac{1}{4Q^2}\right)^{1/2} \omega_0.$$

Equation (15.16) represents damped harmonic motion of frequency ω that has been modified with respect to ω_0 , the frequency of the undamped system ([Fig. 15.5](#)). For $Q = \frac{1}{2}$, the motion is exponentially decreasing with no harmonic component and the system is said to have *critical damping*. For larger values ($Q > \frac{1}{2}$) the motion is a damped harmonic motion. For smaller values ($Q < \frac{1}{2}$) the angular frequency ω is imaginary and the motion is exponentially decreasing.

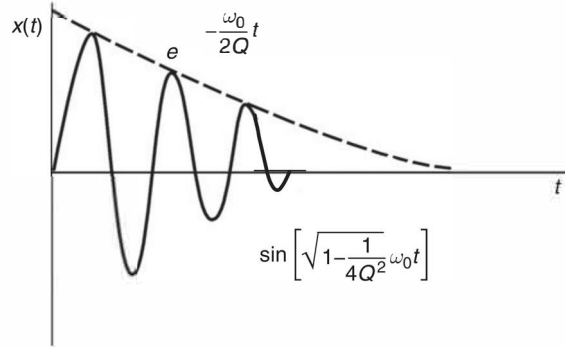


Fig. 15.5 The decrease of amplitude with time in damped harmonic motion.

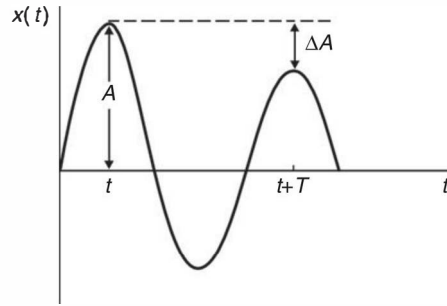


Fig. 15.6 The decrease of amplitude during one period in damped harmonic motion.

In damped harmonic motion, amplitudes decrease exponentially with time. For two times separated by a period $T = 2\pi/\omega$, the ratio of amplitudes according to (15.16) (Fig. 15.6) is given by

$$\frac{x(t)}{x(t+T)} = \exp\left(\frac{\pi}{Q\left(1 - \frac{1}{4Q^2}\right)^{1/2}}\right). \quad (15.17)$$

The logarithm of this quotient is called δ , the *logarithmic decrement*,

$$\delta = \frac{\pi}{Q\left(1 - \frac{1}{4Q^2}\right)^{1/2}} \simeq \frac{\pi}{Q}, \quad Q \gg 1. \quad (15.18)$$

If we consider the decrease in amplitude during one period,

$$\Delta A = x(t) - x(t+T) = A(1 - e^{-\delta}), \quad (15.19)$$

then, for large values of Q , according to (15.18) and (15.19), on taking the first two terms of a series expansion of the exponential function, we obtain

$$\frac{1}{\pi} \frac{\Delta A}{A} = \frac{1}{Q}. \quad (15.20)$$

This expression defines the coefficient Q from the point of view of damped harmonic motion. The factor $1/Q$ represents the ratio of the decrease in amplitude during one period and the initial amplitude. This definition is valid for values of Q that are sufficiently large with respect to unity (the approximation of (15.18)).

In conclusion, we have seen that the inelastic behavior of a material may be represented by the coefficient Q , which is called the *quality factor* and can be defined in various ways. For damped harmonic oscillations, $1/Q$ represents the ratio of the decrease in amplitude during one cycle and its initial value (15.20). This definition of Q can be used to characterize any type of process in media in which the amplitude of motion decreases due to there being a lack of perfect elasticity.

15.2 Wave attenuation

Wave propagation, as we saw in Chapter 5, implies a variation of motion in space and time. Thus, attenuation of wave motion can be considered in time or in space. For a given location wave motion is attenuated with time and, for a given time, it is attenuated with distance. For wave motion, similarly to equation (15.20), we can now define the quality factor as a function of frequency $Q(\omega)$ in terms of energy in the form

$$\frac{1}{Q(\omega)} = \frac{1}{2\pi} \frac{\Delta E}{E}. \quad (15.21)$$

In this definition, $1/Q$ represents the ratio of the elastic energy ΔE dissipated during one cycle of harmonic motion of frequency ω and the maximum or the mean energy E accumulated during the same cycle.

If we consider a harmonic wave of amplitude A that is attenuated so that, after one period or one wave length, its amplitude is $A \exp(-\pi/Q)$, then, since the energy is proportional to the square of the amplitude (Section 5.4), the energy dissipated in one cycle is

$$\Delta E = A^2 \left[1 - \exp\left(-\frac{2\pi}{Q}\right) \right]. \quad (15.22)$$

On taking the ratio $\Delta E/E$, we obtain equation (15.21) in the same way as we derived equation (15.20). Thus for wave propagation we have the quality factor in terms of amplitude (15.20) or of energy (15.21).

Since wave phenomena can be considered as variations in time or space, the energy dissipated during one cycle can be considered as the dissipation during one period or along

one wave length. In this form we define temporal (Q_t) and spatial (Q_s) quality factors. According to (15.20) and (15.21); Q_t represents the wave attenuation with time during one period for a fixed point of space and Q_s represents the attenuation at a given time along a wave length distance.

Wave attenuation can also be expressed by assigning complex values to the frequency and wave number. For a harmonic elastic wave we write,

$$u(x, t) = A \exp[i(k'x - \omega't)], \quad (15.23)$$

where the wave number and frequency are now complex quantities:

$$k' = k + ik^*, \quad (15.24)$$

$$\omega' = \omega - i\omega^*. \quad (15.25)$$

Equation (15.24) becomes

$$u(x, t) = A \exp [i(kx - \omega t) - (k^*x + \omega^*t)]. \quad (15.26)$$

At a fixed time, amplitudes attenuate with distance; and, at a given distance, amplitudes attenuate with time:

$$u(x) = A e^{-k^*x} \cos(kx - \omega t), \quad (15.27)$$

$$u(t) = A e^{-\omega^*t} \cos(kx - \omega t). \quad (15.28)$$

According to definitions of Q_t and Q_s we can easily deduce that

$$\frac{1}{Q_t} = \frac{2\omega^*}{\omega}, \quad (15.29)$$

$$\frac{1}{Q_s} = \frac{2k^*}{k}. \quad (15.30)$$

Since the phase velocity is $c' = \omega'/k'$, it also has a complex value $c' = c + ic^*$. If $\omega^* \ll \omega$ and $k^* \ll k$, the imaginary part of the phase velocity is

$$c^* = c \left(\frac{\omega^*}{\omega} + \frac{k^*}{k} \right). \quad (15.31)$$

In conclusion, Q_t is related to the imaginary part of the frequency and Q_s is related to that of the wave number. A monochromatic wave of frequency ω and phase velocity c travels a distance x in time t and then $k^*x = \omega^*t$. Since $c = x/t$, by using (15.29) and (15.30) we obtain that $Q_t = Q_s$. For non-dispersed waves the spatial and temporal attenuations are equal.

15.3 The attenuation of body and surface waves

15.3.1 Body waves

The attenuation of body waves can be expressed by taking complex values for the velocities of P and S waves, namely, $\alpha' = \alpha + i\alpha^*$ and $\beta' = \beta + i\beta^*$. Since the attenuation of body waves is measured from amplitudes at various distances, the imaginary parts of velocities are related to the spatial quality factor Q_s . For P and S waves we can also define quality factors in a similar way to (15.29) and (15.30):

$$\frac{1}{Q_\alpha} = \frac{2\alpha^*}{\alpha}, \quad (15.32)$$

$$\frac{1}{Q_\beta} = \frac{2\beta^*}{\beta}. \quad (15.33)$$

The complex velocities can now be expressed in terms of the corresponding Q factors:

$$\alpha' = \alpha \left(1 + \frac{i}{2Q_\alpha} \right), \quad (15.34)$$

$$\beta' = \beta \left(1 + \frac{i}{2Q_\beta} \right). \quad (15.35)$$

To introduce the attenuation, we can also consider complex values for the elasticity coefficients μ and K (the rigidity and bulk modulus), $\mu' = \mu + i\mu^*$ and $K' = K + iK^*$, and define the quality factors Q_μ and Q_K :

$$\frac{1}{Q_\mu} = 2 \left(\frac{\mu^*}{\mu} \right)^{1/2}, \quad (15.36)$$

$$\frac{1}{Q_K} = 2 \left(\frac{K^*}{K} \right)^{1/2}. \quad (15.37)$$

According to (4.72) and (4.73), the relations between Q_μ and Q_K and Q_α and Q_β (for values of Q_α and Q_β larger than unity) are

$$\frac{1}{Q_\beta} = \frac{1}{Q_\mu}, \quad (15.38)$$

$$\frac{1}{Q_\alpha} = \frac{4}{3} \left(\frac{\beta}{\alpha} \right)^2 \frac{1}{Q_\mu} + \left[1 - \frac{4}{3} \left(\frac{\beta}{\alpha} \right)^2 \right] \frac{1}{Q_K}. \quad (15.39)$$

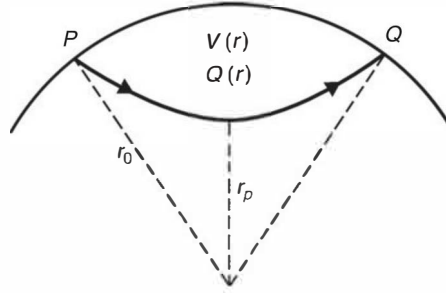


Fig. 15.7

The trajectory of P or S waves inside the Earth with the radial dependence of the velocity $v(r)$ and attenuation $Q(r)$.

In most seismological problems, it is assumed that there is no dissipation of energy in purely compressive or dilational processes and therefore $Q_K = \infty$. Under this hypothesis, from (15.38) and (15.39), we obtain

$$\frac{1}{Q_\alpha} = \frac{4}{3} \left(\frac{\beta}{\alpha} \right)^2 \frac{1}{Q_\beta}. \quad (15.40)$$

If $\sigma = 0.25$, $\alpha = \sqrt{3}\beta$ and the relation gives $Q_\alpha = \frac{9}{4}Q_\beta$. According to (15.31) for a P wave ($c^* = \alpha^*$) we find the relation

$$\frac{1}{Q_\alpha} = \frac{1}{Q_t} + \frac{1}{Q_s}. \quad (15.41)$$

Since for a non-dispersive wave $Q_t = Q_s$, then $1/Q_\alpha = 2/Q_s = 2/Q_t$ and a similar relation for S waves.

In the ray theory approximation for body wave propagation (Chapters 8–10), we are interested in the attenuation along the path of a ray from the focus to the observation point. The attenuation of the amplitude of a monochromatic P wave in the Earth's interior, if at the focus ($t = 0$) the amplitude is A_0 according to (15.28) at a distance s along the ray, is given by

$$A = A_0 e^{-\omega^* t},$$

where t is the travel time. Using (15.29) and that $Q_\alpha = Q_t$, we obtain

$$A = A_0 \exp\left(-\frac{\omega t}{2Q_\alpha}\right) = A_0 e^{-\omega t}, \quad (15.42)$$

where for a homogeneous medium $t = s/\alpha$ (s the length along the ray) and $t^* = t/(2Q_\alpha)$. A similar relation can be written for S waves using the corresponding travel time ($t = s/\beta$) and Q_β . For a spherical Earth of radial symmetry with velocity $v(r)$ and slowness $\eta(r) = r/v(r)$, the quality factor $Q(r)$ depends also on the radius. Using (10.20) for a ray with a surface focus and a ray parameter p , we obtain (Fig. 15.7)

$$t^* = \int_{r_p}^{r_0} \frac{\eta^2(r) dr}{r Q(r) [\eta^2(r) - p^2]^{1/2}}. \quad (15.43)$$

If $\bar{Q}$ is the mean value of $Q(r)$ along the ray and t is the travel time, then an approximation of (15.43) is $t^* = t/(2\bar{Q})$. For the Earth, for epicentral distances between 30 and 90 degrees, t^* is practically constant with values of 1 s for P waves and 4 s for S waves (we will use this approximation on wave form modeling, Section 20.3). This means that S waves attenuate faster than P waves.

Since we do not know the amplitude at the focus, the attenuation is usually determined from amplitude ratios of waves observed at a certain distance between two stations. For body waves, we need observations along similar ray paths, so that the attenuation of the amplitude between the two stations is referred to a certain distance Δs along the ray path inside the Earth. In an approximate form, the attenuation of a body wave of frequency ω along a horizontal distance Δx between two stations may be found from

$$\ln \left(\frac{A_2(\omega)}{A_1(\omega)} \right) = \ln C - \gamma(\omega)\Delta x, \quad (15.44)$$

where $\gamma(\omega)$ is the overall anelastic attenuation of amplitude with horizontal distance and C depends on the geometric spreading (Section 5.9). According to (15.44), for a P wave, $\gamma(\omega) = \omega/(2\bar{\alpha}\bar{Q})$, and $\bar{\alpha}$ and $\bar{Q}$ are average values corresponding to the increment in the horizontal distance Δx . Equation (15.44) illustrates the difference between the influence of geometric spreading and anelastic attenuation on the decrease in amplitude with distance.

15.3.2 Surface waves

The anelastic attenuation of surface waves with horizontal distance (space) and time can be expressed by using the coefficients γ_s and γ_t :

$$u(x) \simeq A \exp \left[-\gamma_s x + i\omega \left(\frac{x}{c} - t \right) \right], \quad (15.45)$$

$$u(t) \simeq A \exp \left[-\gamma_t t + i\omega \left(\frac{x}{c} - t \right) \right]. \quad (15.46)$$

According to equations (15.27)–(15.30), the attenuation coefficients for each frequency, in terms of the quality factors, are given by

$$\gamma_s = \frac{\omega}{2cQ_s}, \quad (15.47)$$

$$\gamma_t = \frac{\omega}{2Q_t}. \quad (15.48)$$

For a dispersed wave, energy propagates with the group velocity U . For amplitudes corresponding to instantaneous frequencies, the phase velocity must be replaced by the group velocity (Section 13.1). If ω_0 is the instantaneous frequency, then, for given corresponding values of x and t , the attenuation in time and space (15.47 and 15.48) is given by

$$A \exp \left(-\frac{\omega_0 t}{2Q_t} \right) = A \exp \left(-\frac{\omega_0 x}{2cQ_s} \right). \quad (15.49)$$

Since the time that the corresponding waves group takes to travel through the medium is $t = x/U$, then we deduce for dispersed waves the relation,

$$\frac{1}{Q_t} = \frac{U}{c} \frac{1}{Q_s}. \quad (15.50)$$

For dispersed waves, the temporal and spatial quality factors are different. In a homogeneous half-space, Rayleigh waves are not dispersed and we can define for them a quality factor Q_R that is related to those of P and S waves:

$$\frac{1}{Q_R} = m \frac{1}{Q_\alpha} + (1 - m) \frac{1}{Q_\beta}, \quad (15.51)$$

$$m = \frac{(2 - b)(1 - b)}{(2 - b)(1 - b) - b/[a(1 - a)(2 - 3b)]}, \quad (15.52)$$

$$a = (c/\alpha)^2, \quad b = (c/\beta)^2,$$

where c is the velocity of Rayleigh waves. In a flat layered medium or one with a depth-dependent velocity distribution, Q values for Rayleigh and Love waves depend on the distributions with depth of $Q_\alpha(z)$, $Q_\beta(z)$, $\alpha(z)$, and $\beta(z)$. Because the depth to which surface waves penetrate depends on their wave length, their Q values are functions of the frequency, $Q_R(\omega)$ and $Q_L(\omega)$.

The attenuation of surface waves can be determined from spectral amplitude ratios of waves of the same frequency at two stations at different distances along the same great circle path relative to the epicenter. For a spherical Earth, according to (12.99) and (15.45),

$$\ln \left(\frac{A_2(\omega)}{A_1(\omega)} \right) = \frac{1}{2} \ln \left(\frac{\sin \Delta_1}{\sin \Delta_2} \right) - \gamma_s(\omega) \Delta x \quad (15.53)$$

where the Δ are angular distances from the epicenter to the stations and Δx is the distance between the stations. The first term on the right represent the geometrical spreading and the second the anelastic attenuation of surface waves along the distance between the two stations for each frequency given by the coefficient $\gamma_s(\omega)$. From $\gamma_s(\omega)$ we can find values of Q_s corresponding to each frequency by using (15.47).

15.4 The attenuation of free oscillations

In Chapter 14, we discussed free oscillations of the Earth considered as a perfectly elastic sphere. If the elasticity is not perfect, the amplitude of each mode of the free oscillations decreases with time due to anelastic attenuation. According to (15.29) and (15.30), the attenuation with time depends on the quality factor Q . For a homogeneous Earth, the attenuation of toroidal modes depends on Q_β and that for spheroidal modes depends on a combination of Q_β and Q_α . For an Earth with radial symmetry, Q_α and Q_β are functions of the radius and Q corresponding to free oscillations varies according to the

eigenfrequencies corresponding to each mode $Q(\omega_n)$. The maximum amplitude for each mode corresponds to a certain depth and thus it is influenced by the values of Q_α and Q_β corresponding to that depth.

We have seen that free oscillations are detected in peaks of the power spectrum corresponding to the frequencies of each mode. In a perfectly elastic medium, the displacements of each mode are harmonic functions of time and their power spectra are delta functions of eigenfrequencies $\delta(\omega - \omega_n)$ (Fig. 15.8a). In the presence of anelasticity, the displacements of each mode are represented by damped harmonic motion:

$$u_n(t) = A_n e^{-\gamma_n t + i\omega_n t}. \quad (15.54)$$

According to (15.30), $\gamma_n = \omega_n/(2Q)$. The power spectrum of (15.54) is

$$|U(\omega)|^2 = \frac{A_n^2}{\gamma_n^2 + (\omega - \omega_n)^2}. \quad (15.55)$$

The spectrum is not a single delta function but it is spread over some frequencies around ω_n . It has a maximum amplitude A_n/γ_n , centered at a frequency ω_n , and a certain width $\Delta\omega$ (Fig. 15.8b). In the frequency domain, Q can be estimated by the spread of the corresponding spectral peak. If $\Delta\omega$ is the width of the spectrum when its amplitude is half its maximum value, we can obtain Q by

$$\frac{1}{Q} = \frac{\Delta\omega}{\omega_n}. \quad (15.56)$$

Since the attenuation increases with $1/Q$, the larger the attenuation the wider the peaks of the spectrum. Broadening of spectral peaks is the main effect of anelasticity on free oscillations. This is a very general effect that applies also to observations in the time domain of propagating pulses in an imperfectly elastic medium. In all cases, the larger the attenuation the wider the pulses become.

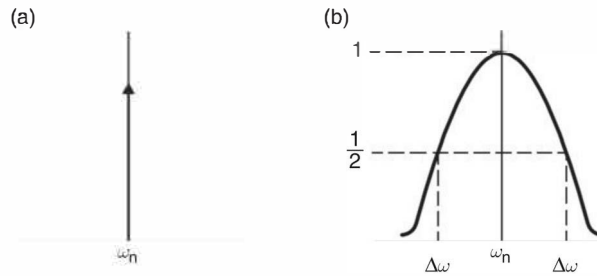


Fig. 15.8

The broadening of the spectral peak of a mode of oscillation due to attenuation.

15.5 The attenuation of coda waves

Waves observed in the last part of a seismogram are called coda waves. This name seems to have been given by Jeffreys in 1929 to waves arriving after surface waves. Today this name is applied to waves of near earthquakes that arrive after Lg waves with amplitudes decreasing exponentially with time (Aki, 1969) (Fig. 15.9). Their observation provides important information about attenuation of waves at short epicentral distances. The attenuation of coda waves has become an important area of research in seismology (Aki and Chouet, 1975; Herraiz and Espinosa, 1987). The attenuation of coda waves is caused by two different effects: anelasticity and scattering of waves due to their interaction with obstacles or heterogeneities in the medium. In the first case, as we have seen, energy is lost by conversion into heat through internal friction (intrinsic attenuation). In the second, energy is distributed through space so that part of it does not arrive at the observation point. Both factors contribute to the attenuation of amplitudes of waves observed at a distance relatively near to the focus. For such short distances, waves propagate mainly through the crust and are affected by its anelasticity and heterogeneity. The total attenuation is given by the coda Q factor, Q_c , which includes both effects:

$$\frac{1}{Q_c} = \frac{1}{Q_i} + \frac{1}{Q_{sc}}. \quad (15.57)$$

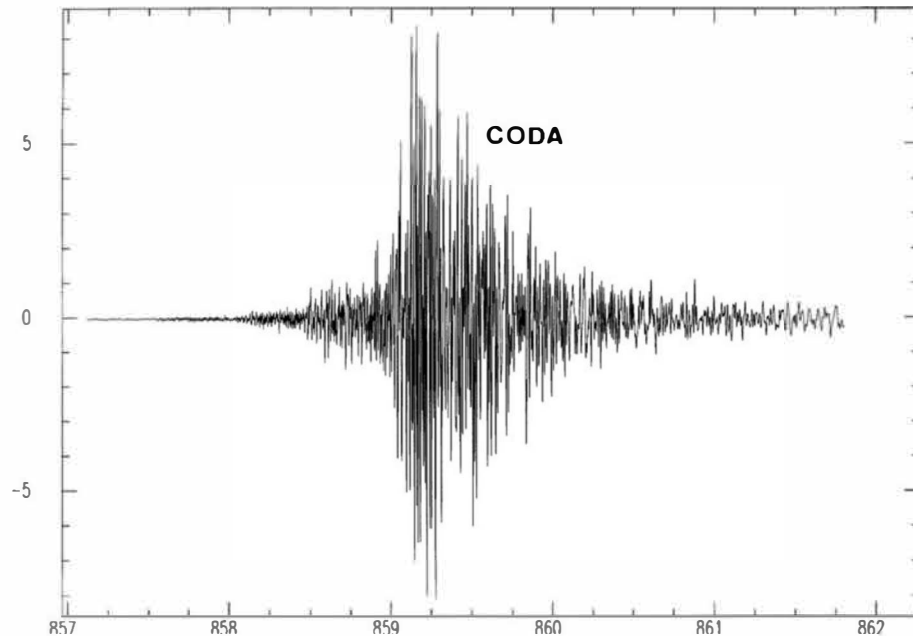


Fig. 15.9

A seismogram of a near earthquake, for which coda waves are shown (the SFUC BB station, for the earthquake of 21 May 1997, $\Delta = 324$ km).

Q_i represents the intrinsic attenuation and is approximately equal to Q_β , indicating that coda waves are principally transverse waves, and Q_{sc} accounts for the attenuation due to scattering phenomena. A formula proposed by Dainty (1981) for Q_{sc} is

$$\frac{1}{Q_{sc}} = \frac{gv}{\omega}, \quad (15.58)$$

where v and ω are the velocity and frequency, and $g = \Delta I/(IL)$ is the dispersion coefficient given by the wave's energy I and the fraction of it ΔI that is lost when it crosses a layer of thickness L where heterogeneities are present. Many other models are used to explain coda wave attenuation. They range from pure diffusion, for which $Q_{sc} = Q_i$, to simple and multiple scattering with complex interaction of waves with obstacles and heterogeneities. In principle, the attenuation of coda waves allows the determination of both anelasticity (Q_i) and heterogeneity (Q_{sc}) of crustal material. However, it is not easy to separate the two effects that contribute to Q_c .

The amplitudes of coda waves attenuate with time in the form,

$$A(\omega, t) = A_0 \exp\left(-\frac{\omega t}{2Q_c}\right). \quad (15.59)$$

Experimentally it has been found that Q_c varies with frequency in the form,

$$Q_c(\omega) = Q_0 \left(\frac{\omega}{\omega_0}\right)^n, \quad (15.60)$$

where Q_0 is the value corresponding to the reference frequency ω_0 . The exponent n varies in the range 0.2–0.4 for high values of Q_0 and is close to unity for low values.

15.6 Attenuation in the Earth

Studies of seismic waves attenuation have led to the determination of the Q distribution in the interior of the Earth. In the first models with radial symmetry, $Q(r)$, the velocity distribution was kept constant and independent from the attenuation. In the formulation of complex values of velocities, (15.34) and (15.35), this implies that the real and imaginary parts are considered independently. More recent determinations have used simultaneous inversion of attenuation and velocities (Anderson and Archambeau, 1964; Lee and Solomon, 1978). The observations used for determination of the attenuation inside the Earth are those of the amplitudes of body, P, S, PcP, ScS, surface, Rayleigh, and Love waves, coda waves, and free oscillations.

In the PREM model, Q_β is considered to be constant in the lower mantle with a value of 320 (Dziewonski and Anderson, 1981). Model SL8 has a variable distribution of Q inside the Earth (Anderson and Hart, 1978) (Fig. 15.10). Its most important features are a lithosphere (0–80 km) with moderate to high values of Q_β in the range 200–500; an upper mantle (80–500 km) with low values of about 110; and a lower mantle (500–2880 km) with a gradual increase from 150 to 500 of Q_β with depth. Q_β is null in the outer core

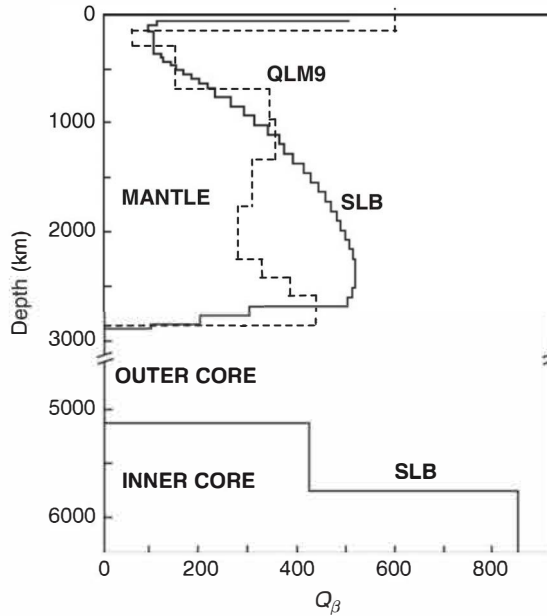


Fig. 15.10

The distribution of Q in the Earth's interior according to model SL8 (Anderson and Hart, 1978) and model QLM9 (Lawrence and Wyession, 2006).

and varies in the range 400–800 in the inner core. Q_K is infinite in the mantle and equal to Q_μ in the inner core. This model is consistent with observations of practically constant values of t^* for distances between 30 and 90 degrees of 1 s for P waves and 4 s for S waves. Other models such as QLM9 (Lawrence and Wyession, 2006) give lower values in the lower mantle, between 300 and 400 with only an increase to 470 above the core (Fig. 15.10). The attenuation in the Earth's interior is principally due to dissipation of energy in shear motion. The absence of energy loss in compressive processes leads to an infinite value of Q_K in the mantle. However, the attenuation of radial modes of free oscillations seems to indicate that it has a large but finite value. Most seismic data can be explained by invoking relatively simple models of the Q distribution with radial symmetry and independence from frequency. For high frequencies, however, there seems to be such a dependence.

For the lithosphere and upper mantle (0–300 km) there are more detailed models of the distribution of Q_β with depth. In the crust, values are relatively low; $Q_\beta \approx 160$. Under the crust (50–100 km) values are higher $Q_\beta \approx 500$. In the asthenosphere or low-velocity layer (100–200 km), values are low, $Q_\beta \approx 125$. The values of Q_c are related to the conditions in the upper crust. In the shallowest layers, there are strong variations in Q_c for different regions with values in the range 120–600. Since Q_c depends on the presence of heterogeneities, low values indicate that the crust is very heterogeneous, which can be associated with seismically active zones, whereas high values can be ascribed to the more homogeneous crust of stable regions. Thus values of Q_c can also be correlated to low or high seismicity.

15.7 Anisotropy

In previous chapters we have always assumed isotropy for the material of the Earth. We have seen that this material is not perfectly elastic; we will see now that it is not perfectly isotropic. Consideration of the lack of isotropy or anisotropy in the Earth is becoming an important subject in seismology and it is related to geodynamic processes. Although deviations from the conditions of isotropy are small, their effect on the propagation of seismic waves can be observed and provides important information (Crampin, 1977; Babuska and Cara, 1991; Gung *et al.*, 2003). Only the most fundamental ideas about wave propagation in anisotropic media are presented.

For an elastic body the relation between stress and strain is given by Hooke's law (4.16):

$$\tau_{ij} = C_{ijkl}e_{kl}. \quad (15.61)$$

As we saw in Section 4.2, C_{ijkl} , the tensor of elasticity coefficients, has 21 independent components. For an isotropic body, these are reduced to two (the Lamé coefficients λ and μ), the tensor is given by (4.18), and Hooke's law by (4.19).

For complete anisotropy without any kind of symmetry, the elastic tensor C_{ijkl} has 21 independent components. If there is some kind of symmetry this number is reduced. For example, there are nine for orthorhombic symmetry, five for hexagonal symmetry, and three for cubic symmetry.

Hexagonal or cylindrical symmetry is often assumed in seismological problems. This symmetry has a principal axis and is called *transverse symmetry*, since any direction normal to this axis has the same properties (Fig. 15.11). According to Love (1927), the five independent elastic coefficients for a medium with hexagonal symmetry with its principal axis in the x_3 direction are

$$\begin{aligned} C_{1111} &= C_{2222} = A \\ &C_{3333} = C \\ C_{3311} &= C_{3322} = F \\ C_{2323} &= C_{1313} = L \\ &C_{1212} = N \\ C_{1122} &= C_{2211} = A - 2N \end{aligned}$$

The strain energy (Section 4.4) is given by

$$\begin{aligned} W &= \frac{1}{2}A(e_{11}^2 + e_{22}^2) + \frac{1}{2}Ce_{33}^2 + F(e_{11} + e_{22})e_{33} + (A - 2N)e_{11}e_{22} \\ &+ \frac{1}{2}L(e_{13}^2 + e_{31}^2) + Ne_{12}^2 \end{aligned} \quad (15.62)$$

On comparing (15.62) with (4.45), we can see that, even for this relatively simple case of anisotropy, the expression becomes more complicated than for the isotropic case. With the Earth, there are several situations that may be treated using this type of anisotropy. For example, in finely stratified media with elastic properties alternating from layer to layer, the material behaves as a whole like an anisotropic medium with hexagonal symmetry, with its

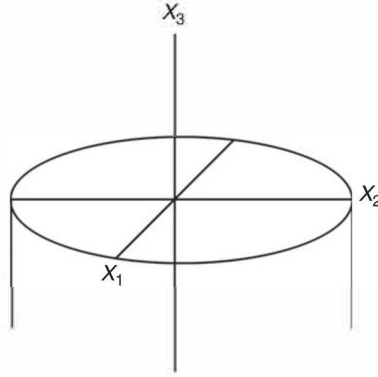


Fig. 15.11 The system of axes in a medium with hexagonal symmetry and the principal axis in the x_3 direction.

principal axis normal to the layers. Since the properties of the medium normal to that axis are the same, this is also called a *transverse isotropic medium*. Another case with the same characteristics is that of a material with cracks aligned in a particular direction that constitutes the principal axis of symmetry.

15.8 Wave propagation in anisotropic media

15.8.1 Body waves

Some properties of wave propagation in anisotropic media can be understood by studying hexagonal symmetry or transverse isotropy. Let us consider the case of the principal axis being in the direction of x_3 . Monochromatic plane waves propagating in the x_3 and x_1 directions are given by

$$u_j = A_j \sin \left[\omega \left(\frac{x_3}{c} - t \right) \right], \quad (15.63)$$

$$u_j = B_j \sin \left[\omega \left(\frac{x_1}{c} - t \right) \right]. \quad (15.64)$$

Using the same procedure as that used in Section 5.6 for an isotropic medium, for propagation in the x_3 direction, we substitute (15.63) into the homogeneous equation of motion (equation (4.67) with $F = 0$) and obtain

$$\left(\frac{1}{\rho} C_{i3k3} - c^2 \delta_{ik} \right) A_j = 0. \quad (15.65)$$

For hexagonal symmetry, on substituting for the components of the elastic tensor in terms of the five independent coefficients A , C , F , L , and N , equation (15.65) is given in matrix form by

$$\begin{bmatrix} L/\rho - c^2 & 0 & 0 \\ 0 & L/\rho - c^2 & 0 \\ 0 & 0 & C/\rho - c^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = 0. \quad (15.66)$$

Then, there are two velocities of wave propagation, given by $(C/\rho)^{1/2}$ and $(L/\rho)^{1/2}$. The first corresponds to waves with only the A_3 component, and, since this is the direction in which waves propagate, it corresponds to P waves with the velocity $\alpha = (C/\rho)^{1/2}$. The second corresponds to waves with components A_1 and A_2 , that is, S waves, and their velocity is $\beta = (L/\rho)^{1/2}$. The displacements are given by

$$u_i^P = (0, 0, A_3) \sin \left[\omega \left(\frac{x_3}{\alpha} - t \right) \right]; \quad \alpha = (C/\rho)^{1/2}, \quad (15.67)$$

$$u_i^S = (A_1, A_2, 0) \sin \left[\omega \left(\frac{x_3}{\beta} - t \right) \right]; \quad \beta = (L/\rho)^{1/2}. \quad (15.68)$$

Wave propagation in this case is similar to that of an isotropic medium.

For propagation in the x_1 direction, on substituting (15.64) into the equation of motion, just like in the previous case, we obtain

$$\left(\frac{1}{\rho} C_{i1k1} - c^2 \delta_{ik} \right) B_j = 0. \quad (15.69)$$

Proceeding as before, we have

$$\begin{bmatrix} A/\rho - c^2 & 0 & 0 \\ 0 & N/\rho - c^2 & 0 \\ 0 & 0 & L/\rho - c^2 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = 0. \quad (15.70)$$

In this case, however, we have three velocities of propagation and therefore three different waves. Since waves propagate in the x_1 direction, B_1 corresponds to P waves and their velocity is $\alpha = (A/\rho)^{1/2}$. The displacements with amplitudes B_2 and B_3 are perpendicular to the direction of propagation and correspond to S waves. There are now two different S waves with different velocities, one with a displacement in the x_2 direction and the velocity $\beta_1 = (N/\rho)^{1/2}$, and another with a displacement in the x_3 direction and the velocity $\beta_2 = (L/\rho)^{1/2}$. The displacements of the three waves are

$$u_i^P = (B_1, 0, 0) \sin \left[\omega \left(\frac{x_1}{\alpha} - t \right) \right]; \quad \alpha = (A/\rho)^{1/2}, \quad (15.71)$$

$$u_i^{S1} = (0, B_2, 0) \sin \left[\omega \left(\frac{x_1}{\beta_1} - t \right) \right]; \quad \beta_1 = (N/\rho)^{1/2}, \quad (15.72)$$

$$u_i^{S2} = (0, 0, B_3) \sin \left[\omega \left(\frac{x_1}{\beta_2} - t \right) \right]; \quad \beta_2 = (L/\rho)^{1/2}. \quad (15.73)$$

By comparison of equations (15.67) and (15.68) with (15.71)–(15.73), we can draw the following conclusions. In an anisotropic medium with hexagonal symmetry, P waves propagate with different velocities along the principal axis of symmetry (x_3) and along

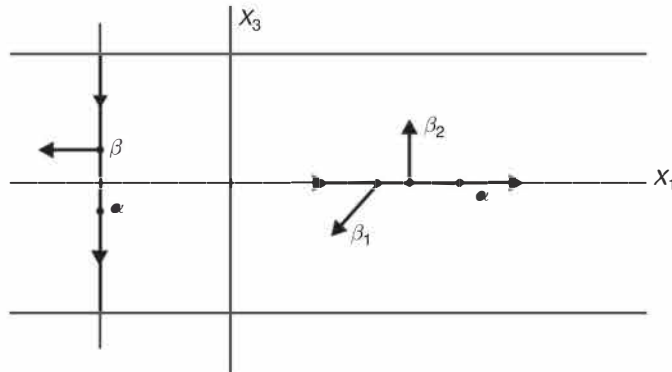


Fig. 15.12

The propagation of $P(\alpha)$ and $S(\beta)$ waves in a medium with hexagonal symmetry and the principal axis in the x_3 direction: (a) waves in the x_3 direction; and (b) waves in the x_1 direction; SH waves have velocity β_1 and SV waves have velocity β_2 .

a direction normal to it (x_1). In the first case there is only one type of S wave and in the second there are two. If x_3 is the vertical direction, then, for waves traveling in the x_1 direction, S1 corresponds to SH and S2 to SV. In this way, SH and SV components propagate with different velocities and are two different waves. This phenomenon is known as *S wave splitting*, since the SV and SH components arrive with a time delay between them (Fig. 15.12). Owing to the type of symmetry, this phenomenon takes place for any orientation of propagation in the plane normal to the principal axis. Thus, for example, in the Earth, S wave splitting occurs during horizontal propagation in media with fine layering of layers with alternating high and low rigidities. In the same medium, P waves propagate with different velocities in the vertical and horizontal directions.

In general, for any type of anisotropy, there are always three types of waves propagating with three different velocities. Choosing the three components of displacement adequately, they are called *quasi-P*, *quasi-SH*, and *quasi-SV* waves. The velocities for these three types of waves change according to the type of symmetry present in the medium. Because of these properties, anisotropy is detected by observations of change in P wave velocity along two perpendicular directions and by observations of S wave splitting. For both effects it is not necessary that the whole medium be anisotropic; only a certain part of it. The time delay between SH and SV components produced in the anisotropic part is preserved, while waves travel along the isotropic part with the same velocity.

15.8.2 Surface waves

The main effect of anisotropy on surface waves is that they can not always be separated into Rayleigh and Love waves, unlike in isotropic media. The radial and transverse components are coupled, forming generalized dispersed surface waves that are a combination of Rayleigh and Love waves. To summarize, we can distinguish three effects of anisotropy on the propagation of surface waves. First, there is a discrepancy in the relation between the phase velocities of Rayleigh and Love waves with respect to that of isotropic media. This

effect is the reason behind the difficulty in finding a unique model of layered isotropic media that agrees with observations of both types of data. Second, there are discrepancies in phase velocities found for trajectories along different azimuths in the same region. Third, there is a departure of the plane of polarization of Rayleigh waves from vertical orientation.

Because of the relatively long wave lengths of surface waves, the type of anisotropy that affects them is produced by oriented heterogeneities in some preferential directions. If heterogeneities are randomly distributed, the medium as a whole behaves isotropically, whereas if they are consistently oriented in a particular direction, they produce an effect of anisotropy. The distinction between heterogeneity and anisotropy is not always an easy one to make and depends on the relation between the dimensions of the heterogeneities and the wave lengths. For anisotropy produced by heterogeneities with some preferential orientation, if the resulting symmetry axis is in a particular horizontal direction, then, along that direction, Rayleigh waves propagate with higher velocities and Love waves with lower velocities than in the isotropic case. For a perpendicular trajectory, the effect is the opposite, with lower velocities for Rayleigh waves and higher for Love waves.

15.9 Anisotropy in the Earth

Observations of the propagation of seismic waves have revealed the presence of anisotropy in the material of the Earth's interior. Essentially two types of anisotropy have been observed. The first has a symmetry with the principal axis in the vertical direction (transverse isotropy) and is due principally to stratifications or horizontal alignments of a structural or mineralogical nature. This results in splitting of S waves with delays in traveling times of SV and SH components and a different relation between the phase velocities for Rayleigh and Love waves depending on their trajectories. The second type is due to preferential alignments of crystals, cracks, or heterogeneities along a particular azimuth (azimuthal anisotropy). This produces effects on the velocities of propagation of waves along trajectories in a particular azimuth in comparison with those perpendicular to them. Some authors distinguish between anisotropy due to stratification and that due to fractures and cracks.

In the shallow part of the crust, anisotropy due to sediment stratification has been observed. A similar effect has also been observed in the lower crust, which has a laminated structure (Section 11.3, Fig. 11.13). Another type of anisotropy observed in the crust is due to the presence of cracks. For a determined stress regime, cracks are oriented in the direction of compression and normal to tensional stresses. Since the crust is subject to tectonic stresses, this may be a very general situation (Crampin, 1978). In subcortical oceanic lithosphere, sea floor spreading may also be responsible for anisotropy. The flow of material from oceanic ridges produces a preferential orientation of olivine crystals along flow lines that is preserved through the aging process, resulting in azimuthal anisotropy.

In the upper mantle, the asthenosphere under the lithosphere is a region of strong anisotropy related to plastic flow of material, that follows the motion of lithospheric plates. Two symmetries may be present, one with a vertical principal axis produced by horizontal

flow that results in higher SH than SV velocities, and another with azimuthal anisotropy along flow lines with higher seismic velocities. For regions with predominantly vertical flow, the opposite effect is observed, with SV velocities higher than SH velocities. These two effects have been observed in surface wave propagation along different trajectories in the Pacific Ocean. Rayleigh waves propagating along trajectories that coincide with plate motion directions have higher velocities than do those with perpendicular trajectories. This anisotropy may extend down to 300 km depth and is mainly associated with the orientation of olivine crystals along flow lines. Variations in velocities for different trajectories may be of magnitude 3–10% (Leveque and Cara, 1985; Nishimura and Forsyth, 1989; Kendall, 2000). For depths below 400 km, observations reveal no appreciable anisotropy and the lower mantle can be considered isotropic. The region of transition between the mantle and the core (the CMB) is considered anisotropic by some authors, but data are not yet sufficient for a definitive conclusion to be drawn.

Already from the PKIKP data accumulated in the early 1980s it appeared evident that a simple laterally homogeneous spherical model for the Earth's inner core couldn't explain the observed travel times (Poupinet *et al.*, 1983). By inverting the mantle-corrected PKIKP travel times, Morelli *et al.* (1986) discovered a marked regional pattern of heterogeneity for waves sampling the innermost part of the inner core. Travel time anomalies of 2–3 s were detected, which were then justified by assuming the presence of a cylindrical anisotropy with a fast velocity direction parallel to the Earth's spinning axis. Such a preliminary hypothesis of a cylindrical anisotropy was further strengthened by normal mode analysis (Woodhouse *et al.*, 1986). These pioneering seismological works set the basis for more complex mineral physics (Mattesini *et al.*, 2013) and geodynamo (Gubbins *et al.*, 2011) scenarios, and a number of mechanisms for the intrinsic elastic anisotropy of the core material have been proposed (Mattesini *et al.*, 2010).

15.10 Summary

In this chapter we have presented two characteristics of the material of the Earth, namely anelasticity and anisotropy, that distinguish it from the perfect elasticity and isotropy heretofore considered. Anelasticity introduces a dependence of the stress with a time derivative of the strain. Its effect on wave propagation determines its amplitude attenuation with time and space, which is expressed by the quality factors Q_t and Q_s . Another way of expressing anelasticity is by taking complex values for angular frequency ω and wave number k and also for P and S wave velocities. Anelasticity also influences free oscillations, broadening the spectral peaks of the modes. Their spectral width $\Delta\omega$ (15.56) is related to the quality factor Q_t for each mode. Another effect of anelasticity is the attenuation of coda waves from near earthquakes. Here, in addition to anelasticity, attenuation is due to the scattering of waves by heterogeneities in the upper crust.

The distribution of the quality factor Q in the interior of the Earth has been found from observations of the attenuation of seismic waves. Models for Q_β show a distribution with high values (200–500) in the lithosphere, low values (125) in the asthenosphere, and

increasing values in the mantle from 150 to 500; at the outer core Q_β is null and has values of 400 and 800 at the inner core. Values of the coda wave attenuation give values for Q_c of between 120 and 600. Low values are related to a very heterogeneous upper crust corresponding to seismically active regions, while high values correspond to a homogeneous stable crust.

Anisotropy also influences wave propagation in the Earth. In seismological problems often a particular case is considered of hexagonal or transverse isotropy. P and S waves are propagated with different velocities along the principal axis of symmetry and in a direction normal to it. In addition to S waves, in the latter case SH and SV components are propagated with different velocities, resulting in splitting of the S waves. Anisotropy affects surface waves so that the relation between Rayleigh and Love wave velocities is different from that for the isotropic case. Anisotropy has been observed in different regions of the Earth's interior. Strong stratification in the sediments and low crust are sources of anisotropy. Another source is the horizontal flow in the asthenosphere related to tectonic plate motion. There is also strong evidence of the presence of anisotropy in the solid material in the inner core.

15.11 Problems

- P15.1 In a harmonic motion the amplitude of a wave of 4 s period is 50 nm and the difference in amplitude in two consecutive maxima is 5 nm. Find the value of Q directly (equation (15.17)) and using the approximation of equation (15.20).
- P15.2 Consider a spherical Earth of radius 5000 km, constant S wave velocity, 6 km/s, and attenuation $Q_\beta = 300$. If the amplitude of an S wave of period of 10 s at the focus is 5 cm, calculate the amplitude at an epicentral distance of 45° and the value of t^* .
- P15.3 If for a P wave of 5 s period and 8 km/s velocity the amplitudes at epicentral distances of 40° and 45° are 54 nm and 45 nm, respectively, without considering the attenuation due to geometrical spreading, determine the average Q value.

Up to this chapter, we have considered the motion produced by earthquakes from the point of view of wave propagation and free oscillations, without considering its origin. Chapters 16 to 20 are dedicated to this subject, that is, the earthquake source, its location, size, and mechanism. As we saw in Chapter 2, most earthquakes are tectonics and caused by sudden motion along faults which releases accumulated tectonic stresses and produces seismic waves. The source of an earthquake can then be considered to be the product of rupturing of a particular fault with a relative displacement of its two sides and release of the accumulated elastic strain that had been produced by tectonic processes. However, as a first approximation, we can consider the source of an earthquake as a point, called its *focus*, from which seismic waves are generated and propagated. The reduction of the focus to a point is called the *point-source* approximation, which is valid if observation points are at a sufficiently large distance compared with source dimensions, and wave lengths are also large. The simplest point source model is the *isotropic point source*, in which the source is considered as a point from which waves propagate with equal amplitude in all directions. In Chapter 2 we have already seen in a descriptive form the problem of the location, time, and size of earthquakes. In this chapter we will quantify this knowledge, defining the focal parameters of earthquakes and their determination, such as location and origin time, the different ways of measuring their size, defining the concepts of intensity and magnitude, and also those of seismic moment and stress drop.

16.1 Location of an earthquake focus

The location of the *point focus* of an earthquake in space and time is given by its geographic coordinates (x_0, y_0), its depth (z_0), and the time of occurrence or *origin time* (t_0). Owing to the dimensions of the actual source, the focal coordinates refer to a certain point, for example, where rupture starts, and the origin time refers to the initiation of the faulting process. The definitions of the parameters of the focal location use this point approximation.

The concept of an earthquake's point focus from which waves propagate in all directions was formalized by Mallet in 1862 in his study of the Neapolitan earthquake of 1857, and it is fundamental to the determination of its location and origin time. The focus at a certain depth is called the *hypocenter* and its horizontal projection the *epicenter*. The hypocentral parameters are the geographic latitude and longitude of the epicenter (x_0 : ϕ and y_0 : λ), the focal depth (z_0 : h), and the origin time (t_0). The desirability of being able to represent earthquakes' foci on maps and the difficulty, sometimes, of determining their depth justifies

common use of the concept of an epicenter. Since the process of rupturing in earthquakes has dimensions and a duration, the point focus is a simplification. We will see that the point focus has different meanings depending on how it is determined.

Lists and catalogs of earthquakes have been published since the seventeenth century ([Chapter 1](#)). The earliest catalogs were based on the report of damage caused, giving approximate information of the location, time, and size. One of the first modern catalogs was published by Mallet in 1858. The first instrumental determinations were started in the beginning of the twentieth century and were published by the Central Bureau of the International Association of Seismology of Strasbourg in 1904. In 1911, the Seismological Committee of the British Association for the Advancement of Science began publication of an earthquake catalog that in 1918 became the International Seismological Summary (ISS). Its continuation, the catalogs of the International Seismological Center (ISC), has been published since 1963. Other agencies such as the United States Coast and Geodetic Survey (USCGS), later the National Earthquake Information Center (NEIC) of the United States Geological Survey (USGS), started to publish hypocenter determinations in about 1930. In many countries there are institutions that perform hypocenter determinations and publish catalogs of regional earthquakes. At present many catalogs are available on websites (see [Chapter 2](#)).

16.1.1 Macroseismic determination of epicenter locations

The first epicenter determinations were based on damage and were started in the second half of the nineteenth century; today they are known as *macroseismic* determinations. They are based on information regarding damage to buildings, fractures and cracks in the ground, and other effects. Michell in 1760 proposed, for the first time, the possibility of using this type of observation. However, the first to perform a true determination of the location and depth of an earthquake's focus was Mallet, in 1862, for the Naples earthquake of 1857. By using the orientation of the lines along which he thought waves had propagated, on the basis of the orientation of cracks and other evidence, he shows in his own words “the successive steps by which finally the depth of the focus whence the impulse that produced the earthquake has been for the first time ascertained and measured in miles and yards” (Mallet, 1862). This is the first clear reference to the concepts of the epicenter and hypocenter and their determination. Early work in determination of foci was done by Milne, Omori, and Mercalli, among many others.

Macroseismic epicenter determination is based on field observation of the effects of an earthquake on the ground and on buildings and other structures. From these observations, intensities are assigned to each location and an intensity or *isoseismal map* is drawn, as we will see later ([Section 16.2](#)). The macroseismic epicenter is located at the central point of the area of maximum intensity ([Fig. 16.1](#)). This may not coincide with the instrumental epicenter, due to site effects that may contribute to the location of maximum damage. The depth of the focus can also be determined from intensity maps, as we will see later, from the distribution of zones of various intensities. As an example, for the Andalusian earthquake of 25 December 1884, from the intensity map of [Fig. 16.1](#), the epicentral location was found to be 36°58'N, 4°00'W and depth between 10 km and 20 km (Udías and Muñoz, 1979). Bakun and Wentworth (1997) have developed a modern method to estimate magnitude and location of earthquakes from intensity data.

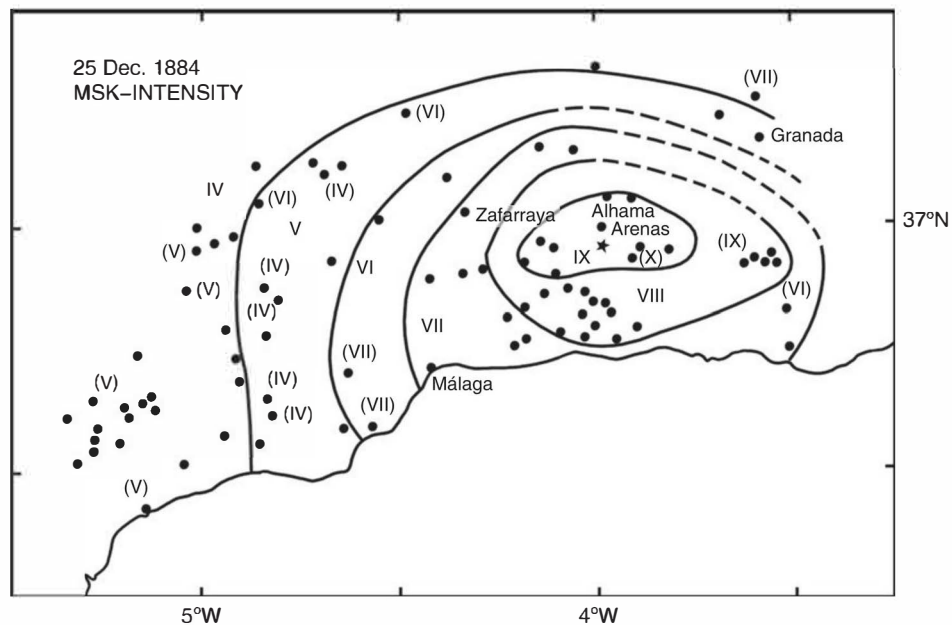


Fig. 16.1

An intensity map of the Andalusian earthquake of 25 December 1884 and location of the macroseismic epicenter (courtesy of D. Muñoz).

Locations from intensity data were used until the installation of seismographic instruments at the end of the nineteenth century. The methods are still used today for determination of the epicenters of *historical earthquakes*, that is, those that occurred before the instrumental period. The study of historical earthquakes is an important part of the evaluation of the seismicity of a region. Large earthquakes may be separated by long periods of time, so it is necessary to extend the study as far into the past as possible. These studies require a search for historical documents and a careful evaluation of the information about damage produced by earthquakes in order to draw reliable intensity maps and estimate their location and size.

The concept of a macroseismic epicenter is not equivalent to instrumental determination based on the times of arrival of seismic waves. An instrumentally determined hypocenter corresponds to the point and time of initiation of the fracture process (when the first waves are generated), whereas a macroseismic epicenter is situated at the central point of the zone of maximum damage. This point may represent, in an approximate way, the projection onto the surface of the central point on the fracture plane. In this sense, it may approximate the horizontal projection of the *centroid* of the fault area, a concept introduced in some methods for the determination of the focal mechanism (Section 20.4).

16.1.2 Instrumental determination. Graphical methods

Instrumental hypocenter determinations are based on the arrival times of seismic waves, mainly P and S waves, recorded on seismograms at stations distributed around the

epicenter. The first methods were graphical, using a map for short distances and a globe for larger or teleseismic distances. Graphical methods determine the location of the epicenter, its origin time, and an estimation of its focal depth. With the generalization of computers for numerical determinations, graphical methods are now no longer used. However, they show the basic ideas behind focal determinations and can be simulated on computer screens. For this reason we present the procedure here.

A first approximation to the location of an earthquake can be made from the seismogram of a single station. From the observed arrival times of P and S waves we can determine the time interval between them ($\delta t = t^S - t^P$), which from travel time curves or tables gives the distance Δ at which the epicenter is located. The amplitudes of the horizontal components of the P wave are in the direction of propagation (Section 5.7), so we can find the azimuth, $a_z = \tan^{-1}(U_{EW}/U_{NS})$. Thus from the distance and the azimuth we can have a first estimation of the location. However, since we do not know from what side the waves come, there is an ambiguity of 180 degrees. From the observed arrival time of P (t^P) and distance Δ , we can determine from the travel time curves the travel time of the P wave t'^P . The origin time t_0 is determined by subtracting the travel time from the arrival time, $t_0 = t^P - t'^P$.

From several stations the procedure is the following: for short distances ($\Delta < 1000$ km), the problem is solved on a map on which are situated the positions of seismographic stations. The method consists of an iterative graphical process of successive approximations. Let us consider N stations ($N > 4$) where we have measured the arrival times of P and S waves of a local earthquake, namely, t_i^P and t_i^S , $i = 1, 2, \dots, N$. It is necessary to have travel time tables or curves for these waves. We start by determining the time interval between arrivals of P and S waves at one or several stations. From the travel time curve, the S–P interval (δt) at one station gives us the distance Δ , and from that we obtain the traveling time of P waves (t'^P) corresponding to that station (Fig. 16.2a) and the origin time t_0 , as shown before. From this value of the origin time, we obtain the travel times for all of the stations $t_i'^P = t_i^P - t_0$ and, from these, using travel time curves or tables, we determine the distances Δ_i for all of the stations. Using these distances (Δ_i) as radii, we draw circles with their centers at the positions of the stations. In the absence of errors, these circles will cross at a point that corresponds to the epicenter, whose coordinates (ϕ_0 and λ_0) are found from the map (Fig. 16.2b). If the circles do not cross at a point, the origin time t_0 is changed a little and the process is repeated until all or most circles nearly cross at one point.

Figure 16.2 shows an ideal example using five stations. In general, all the arcs of circles cannot be made to cross at a single point, but their intersection defines a small area that reflects errors present in determination of arrival time and travel time curves. The epicenter is located at the center point of this area. If the focus is at a certain depth and we have used travel times for a shallow focus, the circles will not cross at a point. In this case, we have to repeat the process with travel times of different depths (e.g. 20, 40, 80 km, etc.) until a good fit is found. In this way, the depth of focus is also determined. We need a minimum of three stations in order to determine the epicenter and four for adjusting its depth. On the website www.sciencecourseware.com/virtualearthquake/ there is an online version of the graphical method. This version uses three stations (vertical components) and travel time curves specific for regional distances (less than 1000 km) for different regions (California, Mexico, Japan, etc.) to estimate the epicenter and local magnitude. For large distances,

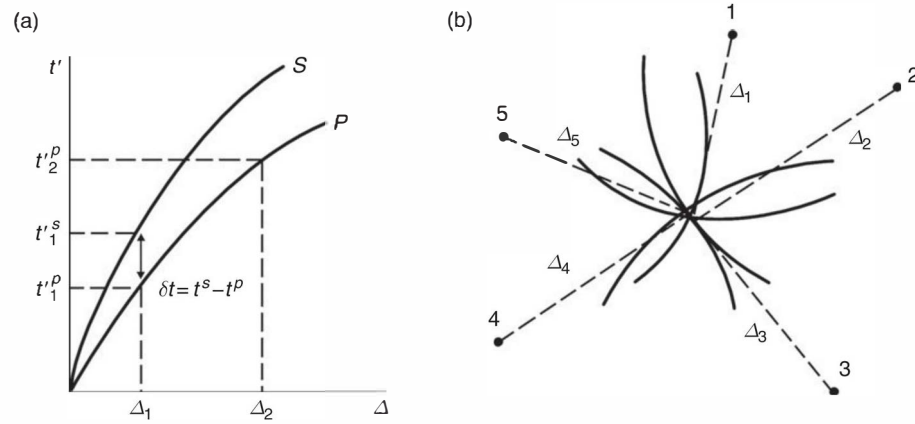


Fig. 16.2

Graphical determination of the origin time and epicenter. (a) Determination of epicentral distance from the S–P time interval ($\delta t = t^S - t^P$) and of the distance Δ_2 from the P wave travel time t_2^P . (b) Location of the epicenter from the intersection of the epicentral distances at five stations.

the method is similar but a terrestrial globe is used for location of the stations. In this case, the depth of the focus can first be estimated, using the time interval between the arrivals of the pP and P rays (Section 11.5). Determination of focal depths by graphical methods is not very exact and, in many cases, they indicate only whether the focus is shallow, intermediate, or deep. In general, if the stations do not surround the epicenter, determinations are not very precise.

16.1.3 Numerical methods

The theoretical basis of numerical methods for hypocenter determination was developed relatively early, with the work of Ludwig Geiger (1910) and Vicente Inglada (1928). Their application, however, was not generalized until the development of digital computers, in around 1960, made fast resolution of the problem for a large number of observations possible. The basis of all methods consists of the linearization of equations representing arrival times and their solution using least squares methods. Various methods have been developed by many authors; we will present only the fundamental ideas that are common to many of them.

As in the graphical methods the basic observations are the arrival times of seismic waves t_i (generally P and S waves, though other phases may also be used) recorded at N stations ($N > 4$) with geographic coordinates ϕ_i and λ_i . The arrival times can be considered as non-linear functions of the coordinates of stations (ϕ_i and λ_i), focal parameters (coordinates, depth, and origin time) (ϕ_0 , λ_0 , h , and t_0), and the distribution of velocities of seismic waves in the Earth's interior. The problem can be linearized by using a Taylor expansion about a set of approximate initial values of the focal parameters (ϕ'_0 , λ'_0 , h'_0 and t'_0) thought to be sufficiently near to the real ones that we can write

$$t_i = t'_i + \delta t + \frac{\partial t_i}{\partial \phi} \delta \phi + \frac{\partial t_i}{\partial \lambda} \delta \lambda + \frac{\partial t_i}{\partial h} \delta h; \quad i = 1, 2, \dots, N, \quad (16.1)$$

where t'_i are the arrival times at each station calculated from the initial solutions (ϕ'_0 , λ'_0 , h'_0 and t'_0) and partial derivatives are evaluated for that solution. We define residuals as the difference between observed and calculated arrival times for each station:

$$r_i = t_i - t'_i; \quad i = 1, 2, \dots, N. \quad (16.2)$$

By substituting (16.2) into (16.1) and expressing the N equations in matrix form, we obtain

$$r_i = A_{ij} \delta x_j; \quad i = 1, 2, \dots, N; \quad j = 1, 2, \dots, 4, \quad (16.3)$$

$$\mathbf{r} = \mathbf{A} \delta \mathbf{x}.$$

The matrix $\mathbf{A}$ ($4 \times N$) is formed by taking the partial derivatives of travel times for each station with respect to the coordinates of the epicenter, depth, and origin time, which are calculated in a numerical form from travel time curves or tables. The four components of the vector $\delta \mathbf{x}$ formed by the increments of the four focal parameters are the four unknowns of the problem. The method is iterative. From the first initial values of focal parameters, by solving equation (16.3), we obtain the first set of increments $\delta \mathbf{x}$ that we add to the initial values to give us new values of the parameters. These new values are used again as initial values and the process is achieved when the increments become small, that is, of the order of errors in observations, or when the overall error has a minimum value. Since the problem is overdetermined, the system (16.3) consists of N equations with four unknowns. For its solution we use a least squares method that minimizes the sum of the squares of the residuals:

$$\varepsilon^2 = \frac{1}{N} \sum_{i=1}^N r_i^2.$$

Several methods may be used for the solution of equation (16.3). Solutions give minimum-error estimates of the parameters and of their standard deviations. A solution can be obtained by multiplying (16.3) from the left-hand side by the transpose of $\mathbf{A}$ and then finding the inverse of the (4×4) matrix $\mathbf{A}^T \mathbf{A}$:

$$\delta \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{r}. \quad (16.4)$$

Another solution can be found by using the generalized inverse matrix (Lanczos, 1957). As we saw in Section 7.4, the matrix $\mathbf{A}$ can be factored as

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T, \quad (16.5)$$

where $\mathbf{\Lambda}$ is a diagonal matrix formed by the square roots of the eigenvalues of $\mathbf{A}^T \mathbf{A}$, $\mathbf{V}$ is a matrix formed by the eigenvectors of $\mathbf{A}^T \mathbf{A}$, and $\mathbf{U}$ is a matrix formed by the eigenvectors of $\mathbf{A} \mathbf{A}^T$. The generalized inverse of $\mathbf{A}$ is given by

$$\mathbf{A}^{-1} = \mathbf{U}^T \mathbf{\Lambda}^{-1} \mathbf{V}. \quad (16.6)$$

Thus the solution of (16.3) is

$$\delta \mathbf{x} = \mathbf{U}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{r}. \quad (16.7)$$

From the matrices $\mathbf{U}$ and $\mathbf{V}$, we can form the covariance matrix $\mathbf{C} = \mathbf{V} \mathbf{\Lambda}^{-2} \mathbf{V}^T$, whose diagonal elements are the variances of the parameters, the resolution matrix $\mathbf{R} = \mathbf{V} \mathbf{V}^T$, whose elements indicate the relative resolution of each parameter, and the information density matrix $\mathbf{D} = \mathbf{U} \mathbf{U}^T$, whose elements indicate which observations contribute the most information to the problem.

If the matrix $\mathbf{A}$ is nearly singular, the problem becomes unstable. One way to avoid this is to introduce an attenuation factor by replacing the matrix $\mathbf{A}^T \mathbf{A}$ in (16.4) by $\mathbf{A}^T \mathbf{A} + k \mathbf{I}$, where k has a small value, before determining the inverse (Marquardt, 1963) and $\mathbf{I}$ the unit matrix. In this form we eliminate the occurrence on the diagonal of elements with values near to zero. In the case of the generalized inverse (16.7), the problem manifests itself in the presence of very small or zero eigenvalues. The matrix $\mathbf{A}$ is replaced by $\mathbf{A} + k \mathbf{I}$, so that eigenvalues near to zero become finite.

Several methods are also used in order to speed up the convergence, such as the introduction of a weighting procedure for observations, use of a *scaling factor*, and *centering*. A scaling factor is introduced so that the columns of the matrix $\mathbf{A}$ have norm unity. Centering consists of replacing the origin time by the weighted mean of the other terms, so that the matrix $\mathbf{A}$ satisfies the condition of being centered. These procedures are commonly used in statistical regression problems. Other new methods use a non-linear method of inversion to estimate the hypocenter and origin time (Lomax, 2005)

The first algorithms and computer programs for hypocenter determination were proposed in around 1960 (Bolt, 1960). At present there are a variety of computer programs or *codes* to perform hypocentral determination. Among those developed for regional distances, widely used is one that is known by the name of HYPO, of which there are several updated versions (Lee and Lahr, 1971). The HYPOINVERSE2000 (Klein, 2000) is a version of HYPO that uses different crustal models to estimate the hypocenter. Crustal models can be multiple to cover different regions, using either flat layers or flat layers with linear velocity gradients. This is a very commonly used program of which there are several versions. Many of these programs may be downloaded from the website <http://earthquake.usgs.gov/research/software>.

16.2 Joint hypocenter determination

Often the problem of hypocenter determination is not well conditioned, even though the number of equations is greater than the number of unknowns. Some of the equations need not be really independent and in many cases the origin time and depth of focus are coupled together. Also, travel times are deduced from average tables or simplified models of the Earth's velocity distribution and do not correspond to real values. For regional distances, heterogeneities of the crust and upper mantle may affect hypocenter determinations. One

solution is to introduce station corrections into the equations, but this is not easy insofar as they depend on ray paths and must be known *a priori*.

A solution for this problem is given by the *joint hypocenter determination* (JHD), that is, the determination of groups of nearby hypocenters using the same group of stations. The method, proposed by Alan Douglas (1967) and developed by James W. Dewey (1972), basically consists in joint determination of hypocenters of M earthquakes using observations from the same set of N stations, introducing as new unknowns N station corrections, which are the same for all events. For simplicity, let us consider that at each station we read only the arrival times of P waves for all shocks. In this case we have NM equations and $4M$ parameters of the M hypocenters. If we add N station corrections, the number of unknowns is $4M + N$. The method uses one earthquake, usually the largest and best-recorded one, as a calibration event for which the solution is supposed to be known previously and the number of unknowns is $4(M - 1) + N$. The linearized equations (15.3) are

$$r_{jk} = A_{sjk}\delta x_{sk} + \delta g_j; j = 1, 2, \dots, N; k = 1, 2, \dots, M; s = 1, 2, \dots, 4, \quad (16.8)$$

where the subindex j refers to stations, k refers to events, and s refers to hypocenter parameters. The δg_j terms are station corrections or anomalies in traveling times for each station. The unknowns are four increments δx_{sk} of the parameters of each event (a total of $4M$) and N corrections δg_j for each station. The method is used for earthquakes that are relatively near to each other, for example, aftershocks or swarms of earthquakes, so that the same station corrections can be used for all shocks. Sometimes this condition may render the problem unstable, since time corrections and traveling times may become linearly dependent. Since the number of equations is greater than the number of unknowns, the problem is solved by a least squares method, just like the normal hypocenter determination.

Among other programs for group determinations, there is, for example, the HypoDD, which uses double-difference (DD) methodology (Waldhauser and Ellsworth, 2000). The DD technique takes advantage of the fact that if the hypocentral separation between two earthquakes is small, compared to the event-station distance and the scale length of velocity heterogeneity, then the ray paths between the source region and a common station are similar along almost the entire ray path, minimizing errors due to the medium.

16.3 Seismic intensity

As mentioned in Section 2.2, the first way to describe the size of an earthquake is in terms of its maximum or epicentral intensity ($I_{\max}$ or I_0). Intensity is a measurement of the shaking of the ground by its effect on people and buildings. The Mercalli–Cancani–Sieberg scale of twelve degrees has served as the basis for later scales. In North America and other countries, the scale commonly used is the modified Mercalli (MM) scale, proposed by Harry O. Wood and Frank Neumann in 1931 and revised by Richter in 1956. In Europe, the most commonly used modern scale was first the MSK scale, published by Serguei Medvedev, Wilhelm Sponheuer, and Vít Kárník in 1967. These two last scales are

Table 16.1 The European Macroseismic Scale 1998 (abridged from Grünthal (ed.), 1998)

Type of structure: masonry (five classes from rubble stone to reinforced brick), Reinforced Concrete (RC) (four classes from RC without Anti-Seismic Design (ASD) to RC with a high level of ASD), and wooden structure.

Vulnerability (classes A–F): A, rubble stone, adobe; B, stone, unreinforced brick; C, brick with RC floors, RC without ASD; D, reinforced brick, RC with minimum ASD; E, RC with moderate ASD; and F, RC with high ASD.

Wooden structures have vulnerabilities from B to F (ranges of vulnerability are given for each type of structure).

Classification of damage: (grades 1–5): 1, negligible to slight (no structural damage); 2, moderate (slight structural, moderate nonstructural); 3, substantial to heavy damage (moderate structural, heavy nonstructural); 4, very heavy (heavy structural, very heavy nonstructural); and 5, destruction (very heavy structural, near or total collapse).

Effects: (a) on humans, (b) on objects and nature, and (c) damage to buildings.

Degrees of intensity

I	Not felt. (a) Not felt.
II	Scarcely felt. (a) Felt by very few.
III	Weak. (a) Felt indoors by a few. (b) Hanging objects swing.
IV	Largely observed. (a) Felt indoors by many, outdoors by a few. (b) Doors and glasses rattle, furniture shakes.
V	Strong. (a) Felt indoors by most, outdoors by a few, strong shaking, people awake. (b) Objects swing, some fall down, doors open or shut, window panes break. (c) Grade 1 damage to a few buildings.
VI	Slightly damaging. (a) Felt by most indoors and many outdoors, many people frightened. (b) Small objects fall, furniture shifts, glassware breaks, animals frightened. (c) Grade 1 damage to many buildings, grade 2 damage to a few.
VII	Damaging. (a) Most people frightened, find it difficult to stand. (b) Furniture shifted and overturned, objects fall, water splashes. (c) Many buildings of class B and a few of C suffer grade 2 damage, many buildings of class A suffer grade 4 damage, especially to their upper parts.
VIII	Heavily damaging. (a) Many find it difficult to stand. (b) Furniture overturned, objects fall, tombstones displaced or overturned, waves seen on soft ground. (c) Many class C buildings suffer grade 2 damage, many class B and a few class C buildings suffer grade 3 damage, many class A and a few class B buildings suffer grade 4 damage, a few class A buildings suffer grade 5 damage, a few class D buildings suffer grade 2 damage.
IX	Destructive. (a) General panic, people thrown to the ground. (b) Many monuments and columns fall or are twisted. Waves seen on soft ground. (c) Many class C buildings suffer grade 3 damage, many class B and a few class C buildings suffer grade 4 damage, many class A and a few class B buildings suffer grade 5 damage, many class D buildings suffer grade 2 damage and a few suffer grade 3 damage, a few class E buildings suffer grade 2 damage.
X	Very destructive. (c) Many class C buildings suffer grade 4 damage, most class A, many class B, and a few class C buildings suffer grade 5 damage, many class D buildings suffer grade 3 damage and a few suffer grade 4 damage, many class E buildings suffer grade 2 damage and a few suffer grade 3 damage, a few class F buildings suffer grade 2 damage.
XI	Devastating. (c) Most class C buildings suffer grade 4 damage, most class B and many class C buildings suffer grade 5 damage, many class D buildings suffer grade 4 damage and a few suffer grade 5 damage, many class E buildings suffer grade 3 damage and a few suffer grade 4 damage, many class F buildings suffer grade 2 damage and a few suffer grade 3 damage.
XII	Completely devastating. (c) Practically all structures above and below ground are destroyed.

practically equivalent. The MSK scale has been updated by the European Macroseismic Scale 1992 (EMS-1992) and more recently by the EMS-1998 (Grünthal (ed.), 1993 and 1998) (Table 16.1). In Japan the scale in use is that of the Japan Meteorological Agency (JMA or Shindo scale) with eight degrees (0–7), created in 1898 and updated in 1996.

The assignment of degrees of intensity from field observations after an earthquake is not free from a certain amount of subjectivity. Although descriptions of degrees of intensity for the scales are well defined, the same situation may be accorded different degrees by different observers.

A modern concept is the *instrumental intensity* (I_I), introduced by Arturo Arias (1970), which is determined from the instrumental observed values of peak ground velocity (PGV) or acceleration (PGA) (Espinosa and Lopez Arroyo, 1977; Atkinson and Kaka, 2007). For example, using PGV instrumental values empirical relations are found for I_I of the type

$$I_I = a + b \log(PGV), \quad (16.9)$$

where the a and b coefficients are estimated for each region.

16.3.1 Isoseismal or intensity maps

Isoseismal (term coined by Mallet) or intensity maps are drawn from observed intensities, with lines separating regions of different degrees on a map (Fig. 16.1). The first such map seems to have been that drawn by Peter N. C. Egen for the earthquake of 1828 in the Netherlands, using his own scale of six degrees (Davison, 1927). Intensity maps, in spite of their lack of precision, are very important tools for establishing the distribution of ground vibration levels due to earthquakes. These maps have a great deal of information regarding the extent and intensity of shaking of the ground, and the response of buildings and other structures. From the point of view of intensity, the size of an earthquake depends not only on its maximum value but also on the extent of areas with various degrees of intensity.

The distribution of intensity on the Earth's surface shown on intensity maps depends not only on the size of an earthquake but also on its focal depth, the attenuation of the shaking of the ground with distance, and local effects. If the epicentral intensity is I_0 , the intensity I , at a certain distance Δ , can be expressed by writing

$$I = I_0 - a \log \left(\frac{1}{h} (\Delta^2 + h^2)^{1/2} \right) - b \left((\Delta^2 + h^2)^{1/2} - h \right), \quad (16.10)$$

where h is the focal depth, a is a coefficient related to geometric spreading (Section 5.9), and b is related to the anelastic attenuation (Section 15.3). Thus, from the intensity map of an earthquake we can estimate its size given by I_{max} or I_0 , the macroseismic epicenter, depth of focus, and values of the coefficients a and b , which give information on how intensities are attenuated in the region near the epicenter (the so-called *near field*) (Fig. 16.3).

In spite of the lack of precision, the information provided by intensity maps is very important and complementary to that provided by the analysis of instrumentally recorded seismic waves. For historical earthquakes, this is the only information available. Even for recent earthquakes, this information is very important, especially from the point of view of

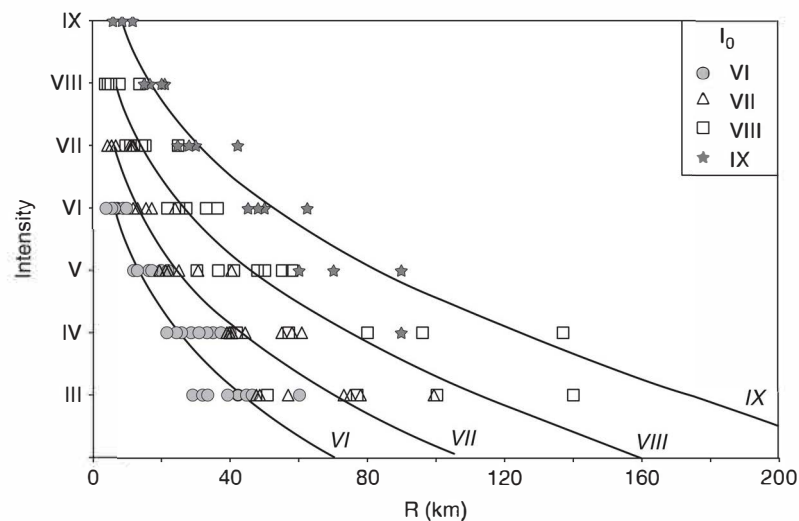


Fig. 16.3 The attenuation of the intensity with distance for earthquakes in the Iberian peninsula (courtesy of C. López Casado).

engineering. The study of earthquakes does not end with the analysis of the shaking of the ground measured by seismographs, but rather extends also to consideration of the damage to buildings and the responses of persons affected by them.

16.4 Magnitude

Intensity, by its very definition, is an indirect measure of the size of an earthquake. A very shallow earthquake can produce very high intensities in a limited region, although its size need not be large. For this reason, the maximum intensity is not always a good indication of the size of an earthquake. Measurement of the size of an earthquake must be done in terms of the energy released at its focus, independently from the damage caused. Since earthquakes are caused by fracturing of crustal material, quantification of their size must represent in some way the energy released by such a phenomenon.

16.4.1 Scales of magnitude

The idea of measuring the size of an earthquake by means of an instrumental estimation of the energy released at the focus led Richter (1935) to the creation of the first scale of magnitude. The concept of magnitude is based on the fact that amplitudes of seismic waves depend on the energy released at the focus, after it has been corrected for their attenuation during propagation (Section 5.4). Using observations of the amplitude of waves of

earthquakes in California at seismographic stations for regional distances ($\Delta < 600$ km), Richter defined the magnitude M in the form

$$M = \log A - \log A_0, \quad (16.11)$$

where A is the maximum amplitude of waves measured on a seismogram in mm (usually Lg waves) and A_0 is a calibration function that depends on distance. Richter defined his scale using records of a particular instrument, the Wood–Anderson torsion seismograph, with amplification 2800 and period 0.85 s (Section 3.1). A_0 corresponds to the amplitude that would be recorded at a given distance for an earthquake of magnitude $M = 0$ (Table 16.2). Calibration of the scale was achieved by assigning the value $M = 3$ to an earthquake that, at a distance of 100 km, is recorded by a Wood–Anderson seismograph with a maximum amplitude of $A = 1$ mm ($\log A_0 = -3$, for $\Delta = 100$ km). This definition is applicable only to surface earthquakes at regional distances and in its original or modified form is today known as the *local magnitude* M_L . Note that for very small earthquakes the magnitude may be negative.

Richter's magnitude for local earthquakes (16.10) can be reformulated in terms of ground motion measured by any type of seismograph with a period near to 1 s by using the formula

$$M_L = \log A + 2.56 \log \Delta - 1.67, \quad (16.12)$$

where A is the maximum amplitude of ground motion (usually corresponding to Lg waves) in micrometers, that is, the amplitude corrected for the instrumental amplification and Δ is the distance in kilometers ($\Delta < 600$ km). IASPEI Magnitude Working Group (2013) (IMWG2013) (Bormann and Dewey, 2014) has reformulated the formulas for magnitude and defined the local magnitude as

$$M_L = \log A + C(R) + D, \quad (16.13)$$

where A is the amplitude in nanometers (nm) and $C(R)$ a function of distance $R < 1000$ km and D is a constant which must be defined for each region. For California the formula is

$$M_L = \log A + 1.11 \log R + 0.00189R - 2.09. \quad (16.14)$$

The extension of the definition of magnitude to earthquakes at large distances (teleseisms) ($\Delta > 600$ km) was done by Gutenberg and Richter between 1936 and 1956. Two scales were defined in terms of ground motion recorded at a distance Δ , one for body waves (generally P waves) and the other for surface waves (Rayleigh waves) (Gutenberg and Richter, 1942, 1956). The first is given by

$$m_b = \log(A/T) + \sigma(\Delta, h), \quad (16.15)$$

where A is the amplitude in micrometers of the ground motion of body waves (corrected for the instrument's response, Chapter 3), T is the period in seconds, and $\sigma(\Delta, h)$ is a calibration term that depends on the distance and focal depth (Table 16.3). Remember that the energy propagated in a wave is proportional to the square of A/T (Section 5.4, eq. (5.51)). Generally, A is measured as the maximum amplitude of the P wave group on vertical component seismograms of short periods ($T \approx 1$ s) instruments. IMWG2013 has reformulated the formula as

Table 16.2 The calibration term in Richter's magnitude.

Δ (km)	$-\log A_0$
10	1.5
20	1.7
30	2.1
40	2.4
50	2.6
100	3.0
150	3.3
200	3.5
300	4.0
400	4.5
500	4.7
600	4.9

Table 16.3 The calibration term of the magnitude m_b (P_z) for a shallow shock.

Δ (degree)	Σ
20	6.00
25	6.45
30	6.65
35	6.70
40	6.70
45	6.80
50	6.85
60	6.90
70	6.95
80	6.90
90	7.00
100	7.40

$$m_b = \log\left(\frac{A}{T}\right) + Q(\Delta, h) - 3.0, \quad (16.16)$$

where A is given in nm and $Q(\Delta, h)$ is equal to $\sigma(\Delta, h)$. For digital broad-band instruments (Section 3.5), the formula is

$$m_{b-BB} = \log\left(\frac{V_{\max}}{2\pi}\right) + Q(\Delta, h) - 3.0, \quad (16.17)$$

where $V_{\max}$ (the peak ground velocity, PGV) is given in nm/s.

For surface waves, the magnitude scale is valid only for surface earthquakes at distances greater than 15° . The formula has the general form

$$M_S = \log(A/T) + \alpha \log \Delta + \beta, \quad (16.18)$$

where A is the maximum amplitude in micrometers of the ground motion of Rayleigh waves, T is the period in seconds (approximately 20 s), Δ is the distance from the epicenter in degrees, and α and β are two calibration constants. The original values given to these constants by Gutenberg (1945) are $\alpha = 1.656$ and $\beta = 1.818$. In 1964, the IASPEI adopted the values $\alpha = 1.66$ and $\beta = 3.3$ (Vaněk *et al.*, 1962).

The IMWG2013 formula with amplitudes in nm is

$$M_{s_20} = \log\left(\frac{A}{T}\right) + 1.66 \log \Delta + 0.3, \quad (16.19)$$

and for digital broad-band instruments with $V_{\max}$ in nm/s,

$$M_{s_BB} = \log\left(\frac{V_{\max}}{2\pi}\right) + 1.66 \log \Delta + 0.3. \quad (16.20)$$

There are other magnitude scales based on amplitudes of certain waves. One introduced by Otto Nuttli (1974), which is valid for regional distances, uses the maximum amplitude of Lg waves (Section 13.7). These waves are the maximum recorded amplitudes on short-period seismograms for continental paths at regional distances:

$$M_{Lg} = \log A + 0.83 \log \Delta + 0.434 \gamma (\Delta - 0.09) + 3.81, \quad (16.21)$$

where A is the amplitude of the ground motion of Lg waves in micrometers, Δ is the distance in degrees, and γ is an attenuation coefficient that is different for each region (for the central USA, $\gamma = 0.07$; for California, $\gamma = 0.53$). IMWG2013 gives the formula,

$$m_b Lg = \log A + 0.833 \log R + 0.4343 \gamma (R - 10) - 0.87, \quad (16.22)$$

where A is in nm, R in km, and the attenuation γ must be defined for each region.

For near earthquakes ($\Delta < 200$ km), owing to the large magnification of modern seismographs, even relatively small earthquakes saturate records and maximum amplitudes cannot be measured. This situation has led to a scale of magnitude based on the duration of a seismic signal instead of its amplitude. The first attempt to use the duration to determine magnitudes was by Ede Bisztricsány (1958). The magnitudes for local earthquakes based on durations have formulas such as

$$M_\tau = a \log \tau - b + c \Delta, \quad (16.23)$$

where τ is the duration of the earthquake signal in seconds and the constants a , b , and c are adjusted for each region so that values of M_τ correspond to those of M_L . For California these constants are $a = 2.2$, $b = 0.87$, and $c = 0.0035$ (Lee *et al.*, 1972).

A different type of magnitude scale was introduced by Hiroo Kanamori (1977) in order to avoid the problem of saturation that afflicts all other scales (Hanks and Kanamori, 1979).

This scale is based on the determination of the seismic moment and is called the *moment magnitude* scale:

$$M_w = 0.67 \log M_0 - 6.07, \quad (16.24)$$

where M_0 is the *scalar seismic moment*, given in Nm, that will be defined later (Section 16.6) and is determined from amplitude spectra at low frequencies or observations of fault areas and amount of slippage.

For earthquakes for which there are no instrumental records, but only macroseismic information concerning the damage they caused (historical earthquakes), magnitudes can be estimated from the epicentral intensity I_0 . An early relation proposed by Sponheuer (1960) (h depth in km) is

$$M = 0.661 I_0 + 1.7 \log h - 1.4. \quad (16.25)$$

Other relations have been developed for specific regions, for example, for Central and Eastern United States (Boyd and Cramer, 2014),

$$M = (I_0 - 2.89 + 0.00277 \Delta + 2.10 \log \Delta) / 1.36 \quad (16.26)$$

$$(\Delta = \text{distance}),$$

and for Europe (Gutdeutsch *et al.*, 2002),

$$M = 0.684 I_0 + 1.559 \log h - 1.514. \quad (16.27)$$

16.4.2 The saturation of magnitude scales

Most magnitude scales depend on the frequency of the waves used for their determination. For this reason it is not possible to define a single scale that is valid for the whole range of observed magnitudes (approximately from -1 to 9). Since the definition of the two teleseismic scales m_b and M_s , it has been observed that they coincide only for values of about 6.5 . For smaller magnitudes m_b is larger and for greater ones M_s is larger (Fig. 16.4). The relation between the two magnitudes established by Gutenberg and Richter (1956) is

$$m_b = 0.63 M_s + 2.5. \quad (16.28)$$

This indicates that the size of small earthquakes ($M < 6.5$) is better measured by m_b and that of large ones ($M > 6.5$) by M_s . This is an example of the saturation of the scales. The scale for m_b becomes saturated at about 6.5 and larger earthquakes do not give greater values. The scale for M_s which underestimates the size of small earthquakes ($M < 6.5$) behaves well for those in the range 6.5 – 8 , but saturates above that value.

This phenomenon is due to the fact that the amplitude spectrum is displaced toward low frequencies with increasing size of earthquakes.

As we have mentioned, the M_s scale is saturated at about $M = 8$ and the size of very large earthquakes is not measured well. These earthquakes produce fractures hundreds of kilometers long with displacements of several meters and waves of 20 s are not representative of the

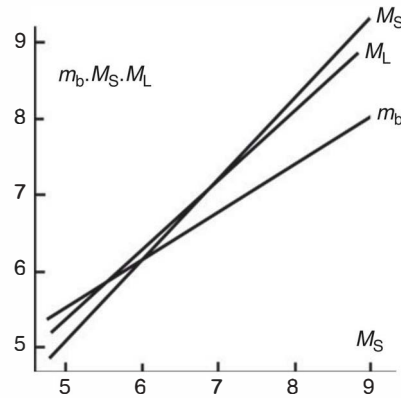


Fig. 16.4 The relation among M_S , m_b and M_L .

energy radiated. This problem is solved with Kanamori's moment magnitude M_W (16.23). This scale does not depend on frequency and can be used for the whole range of size of earthquakes, from very small to very large, up to values of about 9.5. However, its determination is not as simple as those of other scales (direct measurement of an amplitude or duration), since it requires determination of the seismic moment from spectra of seismic waves or other methods.

16.5 Seismic energy

The first references to the energy produced by an earthquake are among others those of Francis Reid, Galitzin, and Manuel Sánchez Navarro-Neumann between 1911 and 1916, with their estimates for the energies of some large earthquakes. The energy propagated by seismic waves is proportional to the square of their amplitude (Section 5.4) and, thus, the magnitude is proportional to the logarithm of the energy. Gutenberg and Richter (1942, 1956) established the first empirical relations between the magnitude and the energy:

$$\log E_S = 2.4m_b - 1.2, \quad (16.29)$$

$$\log E_S = 1.5M_S + 4.8, \quad (16.30)$$

where E_S in joules is the energy propagated in seismic waves, which is often called the *seismic energy*. According to (16.29), an earthquake of $M_S = 8$ has a seismic energy of 10^{17} J. For comparison, a nuclear explosion of 5 megatons (Amchitka, Alaska, 1971) has an energy of 10^{15} J, and is equivalent to an earthquake of magnitude 6.7. If we calculate the approximate energy of all earthquakes that happen in one year we obtain a value in the range 10^{18} – 10^{19} J. About 90% of this energy corresponds to earthquakes with magnitudes equal to and larger than 7. This energy is approximately equal to the global energy consumption in one year.

The total energy E released by an earthquake is the sum of the seismic energy E_S and the energy E_R dissipated by inelastic phenomena and in the form of heat at the focus:

$$E = E_S + E_R. \quad (16.31)$$

Since the only part of E we can measure is the seismic energy, we can consider this as a fraction of the total energy by using the *seismic efficiency coefficient* η ,

$$E_S = \eta E. \quad (16.32)$$

This coefficient has a value less than unity, which is difficult to estimate since we cannot measure the total energy released by earthquakes. For nuclear explosions whose yields are known, values obtained for η vary from case to case according to the conditions of the medium in which they are produced. The mechanism of explosions is very different from that of earthquakes, so the results are not comparable.

16.6 The seismic moment, stress drop, and average stress

The magnitude of an earthquake is related to the energy released in the form of seismic waves and is independent of the mechanism of its generation. Another measure of the size of an earthquake is the seismic moment M_0 , which was introduced by Aki (1966). It is based on the idea that earthquakes are caused by breaking of faults or shear fractures in the Earth's crust, as we saw in Section 2.3, and defined as

$$M_0 = \mu \overline{\Delta u} S, \quad (16.33)$$

where μ is the shear or rigidity modulus, $\overline{\Delta u}$ is the mean value of the slip or displacement on the fault plane, and S is the area of the fault plane. In cgs units the seismic moment is given in dyn cm and in SI units it is in Nm. The seismic moment includes the area of the fault, slip, and strength of the material, and thus constitutes a good physical measure of the size of an earthquake. Its value can be obtained from the spectral amplitudes of seismic waves at low frequency (Section 20.5) or from the field estimations of the slip and area of the faulting after an earthquake.

In a simplified model of fracture, the relative slip Δu of the two sides of a fault is due to the shear stress acting, which, at a given moment, exceeds the strength of the material or the friction that maintains the fault locked. If the shear stresses acting on the fault plane before and after an earthquake are σ_0 and σ_f , we can define two new parameters, namely, the *average stress* $\bar{\sigma}$ (the mean value of the stresses acting before and after the earthquake) and the *stress drop* $\Delta\sigma$ (the difference between them) (Fig. 16.5):

$$\bar{\sigma} = \frac{1}{2}(\sigma_0 + \sigma_f), \quad (16.34)$$

$$\Delta\sigma = \sigma_0 - \sigma_f. \quad (16.35)$$

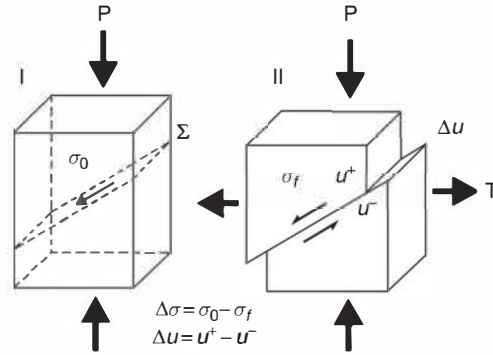


Fig. 16.5

The stress acting before (I) and after (II) the occurrence of a shear fracture of slip Δu and drop in stress $\Delta\sigma$.

The drop in stress represents the part of the stress acting that is employed in producing the slip Δu of the fault. Since σ_0 and σ_f are static initial and final values, $\Delta\sigma$ is called the *static stress drop*. If $\sigma_f = 0$, the stress drop is total and $\Delta\sigma = 2\bar{\sigma}$. Owing to the friction between the two sides of a fault there is always some residual stress σ_f after the fracturing has finished. The initial stress σ_0 is the tectonic stress responsible for the strain in the focal region.

In a simplified form, the total release of energy during fracturing can be expressed by

$$E = \bar{\sigma} \Delta u S \quad (16.36)$$

($\bar{\sigma}S$ represents a force). On substituting into this the seismic moment (16.32), we obtain

$$E = \frac{\bar{\sigma}}{\mu} M_0. \quad (16.37)$$

If the stress drop is total, equation (16.37) gives

$$E = \frac{\Delta\sigma}{2\mu} M_0. \quad (16.38)$$

This expression relates the total energy released by an earthquake to the seismic moment and the stress drop. For a shear fracture, the stress drop is proportional to the slip of the fault, $\Delta\sigma = \Delta u/L'$, where L' is a length dimension of the fault plane (for example, for a circular fault $L' = a$, the radius, whereas for rectangular faults $L' = D$, the width). The stress drop is then given by

$$\Delta\sigma = C\mu \frac{\Delta u}{L'}, \quad (16.39)$$

where C is an adimensional factor that depends on the shape of the fracture (e.g. $C = 7\pi/16$ for a circular fault). For a circular fault, by substituting (16.39) into (16.33), we obtain the relation between the seismic moment and the stress drop:

$$M_0 = \frac{16}{7} a^3 \Delta\sigma. \quad (16.40)$$

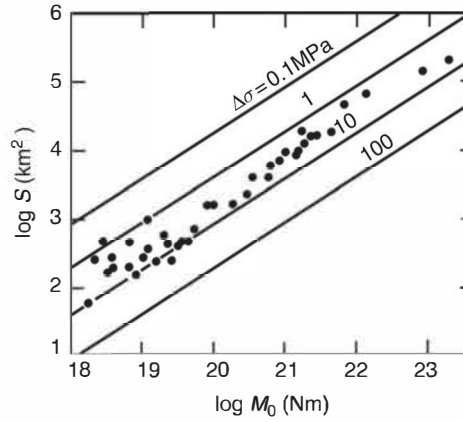


Fig. 16.6

The relation between the fault area S and the seismic moment M_0 with lines of equal drop in stress $\Delta\sigma$ (modified from Kanamori and Anderson (1975) (with permission from the Seismological Society of America)).

Then, if we know the seismic moment and the dimensions of the fracture, we can determine the stress drop:

$$\Delta\sigma = \frac{7}{16a^3} M_0. \quad (16.41)$$

Since the fault radius is present in this equation raised to the power of three, small errors in its estimate will produce large errors in determinations of the stress drop. If we substitute the fault area ($S = \pi a^2$) into (16.41), we obtain

$$M_0 = \frac{16\Delta\sigma}{7\pi^{3/2}} S^{3/2}, \quad (16.42)$$

and, on taking logarithms,

$$\log M_0 = \frac{3}{2} \log S + \log \left(\frac{16\Delta\sigma}{7\pi^{3/2}} \right). \quad (16.43)$$

From this equation it follows that, if the stress drop is constant for all earthquakes, then $\log S$ is proportional to $\frac{2}{3} \log M_0$. That this hypothesis is valid for a large range of magnitudes has been shown empirically (Kanamori and Anderson, 1975). For earthquakes of moderate and large magnitudes ($M > 5$), $\Delta\sigma$ has values in the range 1–10 MPa (10 to 100 bars) and a mean value of 6 MPa (60 bars) (Fig. 16.6). Kanamori and Anderson (1975) suggested that earthquakes that take place at plate boundaries (*interplate shocks*) have lower stress drops (of 3 MPa (30 bars)) than do those in plate interiors (*intraplate shocks*), for which stress drops are about 10 MPa (100 bars). The mean stress drop (6 MPa) is of the same order of magnitude as the value suggested by Chûji Tsuboi (1956) for the critical strain of the Earth's crust.

Constancy of the stress drop is required in the definition of Kanamori's moment magnitude M_w . The formula for the moment magnitude (16.22) is obtained by substitution

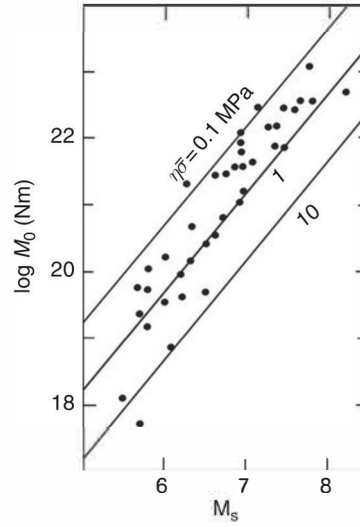


Fig. 16.7

The relation between the seismic moment M_0 and the surface wave magnitude M_S with lines of equal apparent average stress $\eta\bar{\sigma}$ (modified from Kanamori and Anderson (1975)) (with permission from the Seismological Society of America).

of (16.37) into (16.29), assuming a constant value for $\Delta\sigma/\mu = 10^{-4}$ and solving for M_S , which is now renamed M_W . The moment magnitude is, then, the magnitude derived from the seismic moment, under the hypothesis of a constant stress drop that satisfies the relation of Gutenberg and Richter between the surface wave magnitude and the energy (16.29).

By substituting for the energy in terms of the seismic energy (16.31) we can define from (16.36) the *apparent average stress*,

$$\eta\bar{\sigma} = \mu \frac{E_S}{M_0}. \quad (16.44)$$

A simple way to relate the magnitude M_S to the seismic moment M_0 is by using equations (16.30) and (16.44), resulting in

$$\log M_0 = \frac{3}{2} M_S + 11.8 - \log \left(\frac{\eta\bar{\sigma}}{\mu} \right). \quad (16.45)$$

If $\eta\bar{\sigma}$ is constant, we have a linear relation between $\log M_0$ and M_S with a slope equal to $\frac{3}{2}$. Observations agree with this hypothesis, although there is a certain dispersion in the data, especially for very large earthquakes, that may be due to saturation of the M_S scale (Kanamori and Anderson, 1975) (Fig. 16.7). In this case, there is also some evidence that mean stresses for intraplate earthquakes (5 MPa) are larger than those for interplate earthquakes (1.5 MPa).

If the drop in stress is total, we can make the approximation that $\eta\bar{\sigma} \approx \Delta\sigma/2$. According to Egon Orowan's fracture model, the residual energy E_R corresponds to the energy lost by friction and, similarly to (15.26), it is given (Orowan, 1960) by

$$E_R = \sigma_{fr} \overline{\Delta u} S, \quad (16.46)$$

where σ_{fr} is the dynamic frictional stress during fracturing. The situation for which $\sigma_f = \sigma_{fr}$, known as *Orowan's condition*, implies that the stress drop is total and thence $\eta\bar{\sigma} \approx \Delta\sigma/2$. This means that the stress that does not contribute to the slip is lost in friction as heat. When this condition is satisfied, the energy given by (16.37) is the minimum estimation of the total energy E . If the stress drop is not complete, the final stress can be greater or less than the friction stress and the average stress can be greater or less than half the stress drop.

In conclusion, the seismic moment M_0 can be considered the best physical measure of the size of an earthquake. This quantity assumes the mechanism of shear fracture. This model also introduces the concepts of the stress drop and average stress. Through the relation of these quantities to the energy released during fracturing and lost as heat, we have found useful relations among them.

16.7 Summary

The first approximation to the source of an earthquake is that of a point focus from which seismic waves are propagated. Its location in space and time is given by the geographical coordinates, latitude and longitude, and its depth and the origin time (ϕ, λ, h, t_0). The size of an earthquake is given by two concepts: intensity and magnitude. Intensity is based on the damage produced at different points and magnitude is a measure of the energy released at the focus, obtained from the amplitude of the waves recorded by seismograms.

For earthquakes before the instrumental period (historical earthquakes) estimates of the location and size can be obtained from the distribution of intensity and its maximum value (macroseismic location and I_0). For the instrumental period the location and origin time are determined from the time of arrival of seismic waves (P, S, and other phases). At present the determination is made using numerical methods and computer programs. A special method is that of joint hypocentral determination for groups of nearby earthquakes. Magnitude is defined in different ways, for example, from local recordings at near stations (M_L), from teleseismic body waves (m_b), from surface waves (M_s), and from the determination of the scalar seismic moment (M_w). From magnitude we can estimate the energy released at the focus.

From consideration of the source of an earthquake as the rupture on a fault, other parameters are defined. The scalar seismic moment (M_0) is the product of the rigidity at the focal region, the fault area, and the average slip, and can be obtained from the spectral amplitudes at low frequency of seismic waves. From this value we can find the moment magnitude M_w . The static stress drop $\Delta\sigma$ represents the difference in the stress acting at the fault before and after the earthquake and the average stress $\bar{\sigma}$ its average value during the faulting process. The relations between these parameters and energy and magnitude have been derived.

16.8 Problems

P16.1 Consider three stations of coordinates:

St1 = 36.2 °N, 4.8 °E; St2 = 37.0 °N, 2.4 °E; St3 = 38.6 °N, 4.0 °E.

An earthquake is recorded at the three stations with the following respective S–P time intervals:

$$t_1^{S-P} = 26.7 \text{ s}$$

$$t_2^{S-P} = 27.0 \text{ s}$$

$$t_3^{S-P} = 22.5 \text{ s}.$$

Given that the focus is at 60 km depth, the crust is formed by a layer with 65 km thickness, constant P wave velocity equal to 6 km/s, and Poisson ratio 1/4, calculate the coordinates of the epicenter.

P16.2 Using the seismogram of the May 12, 2015, Nepal earthquake recorded at the UCM station (Madrid, Spain) (EM_3_1) estimate the m_b and M_s magnitudes. Use that 1 count $\approx$ 1 nm.

P16.3 Using the value of M_s obtained in problem P16.2 and assuming that $M_w = M_s$ and the seismic efficiency coefficient $\eta = 0.7$, calculate the total energy, the stress drop, and the apparent stress. The rectangular fault dimensions are 60 km by 20 km and the shear modulus, 4.41×10^4 MPa.

In the [previous chapter](#) we have considered a simple isotropic point model for the earthquake source. This model was enough to determine its origin time, location, and size. We have already mentioned in [Chapter 2](#), in a descriptive way, that most earthquakes are caused by breaking on faults due to acting tectonic stresses. In this chapter we will present basic ideas about the source mechanism. The problem with the source mechanism is in relating observed seismic waves to the parameters that describe the source. In the direct problem, theoretical seismic wave displacements are determined from source models, whereas in the inverse problem, the parameters of source models are derived from observed wave displacements. The first step in both problems is to define the seismic source in terms of a mechanical model that represents the physical fracture and derive from this the elastic wave displacement field. These models or representations of the source are defined by parameters whose number depends on their complexity. Simple models are defined by a few parameters, whereas more complex ones require a larger number of parameters (Madariaga, 1983; Udías, 1991; Koyama, 1997; Udías *et al.*, 2014). We begin with formulation in terms of equivalent forces to proceed to those of fractures and dislocations. For mathematical formulation of the problem, a Green's function for an infinite medium is presented and applied to the problem of radiation of seismic waves from a point source shear dislocation.

17.1 Earthquakes and faults

The causes of earthquakes have interested man since antiquity. As was mentioned in [Section 1.1](#), various ideas have been proposed from the time of the ancient Greek natural philosophers up to the present day. During the nineteenth century, systematic field studies were begun following earthquakes and the first attempts to relate these to tectonic processes were made by Mallet (Naples, Italy, 1857), Koto (Neo, Japan, 1891), and Oldham (Assan, India, 1897), among others. With increasing numbers of field observations and improving precision of localization of epicenters, the correlation between earthquakes and faults became clearer. Authors such as Suess, Koto, Montessus de Ballore, and Sieberg assigned the cause of earthquakes to stress accumulated in the Earth's crust by tectonic processes and their release by its fracture. The first mechanical model was presented by Francis Reid (1911) in order to explain the origin of the San Francisco earthquake of 1906. His theory, known as *elastic rebound*, proposes that earthquakes take place by fracturing of the Earth's crust in faults, with total or partial release of the elastic strain accumulated in a region owing

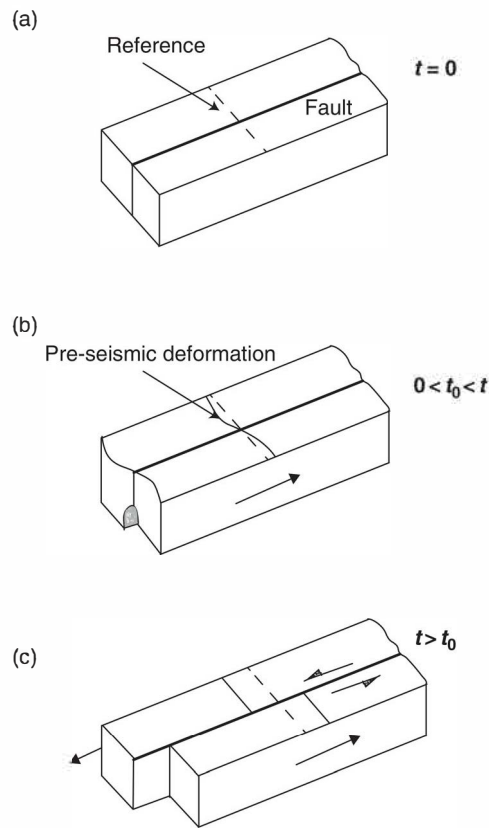


Fig. 17.1

Schematic representation of the elastic rebound theory: (a) initial situation, $t = 0$; (b) pre-seismic deformation due to applied stress; (c) displacement after earthquake occurrence at time t_0 .

to tectonic stress (Fig. 17.1). According to plate tectonics, which was developed in the 1960s, tectonic stresses are ultimately related to the relative motion of lithospheric plates.

An earthquake can, then, be considered to be produced by rupturing of part of the Earth's crust in a fault, with a relative displacement of its two sides, and the release of accumulated elastic strain that had been produced by tectonic processes. Fracture takes place in the brittle part of the crust (about 20 km thickness) called the *seismogenic layer*. The place from where earthquakes originate is called the *focal region*. The parameters that define its mechanism are those that describe the motion of a fracture or fault. These are the following (Fig. 17.2): The *azimuth*, ϕ , is the angle between the trace of the fault (the intersection of the fault plane with the horizontal) and North ($0^\circ \leq \phi \leq 360^\circ$); the angle is measured so that the fault plane dips to the right-hand side. The *dip*, δ , is the angle between the fault plane and the horizontal at a right angle to the trace ($0^\circ \leq \delta \leq 90^\circ$). The *slip angle* or *rake*, λ , is the angle between the direction of relative displacement or slip and the horizontal measured on the fault plane ($-180^\circ \leq \lambda \leq 180^\circ$); λ is negative for *normal faults* and positive for *reverse faults*.

According to the values of δ and λ , we have different types of faults, namely, for $\delta = 90^\circ$ and $\lambda = 0^\circ$, *strike slip* faults; for $\delta = 90^\circ$ and $\lambda = 90^\circ$, *vertical* faults; for $\delta > 0^\circ$, *inclined*

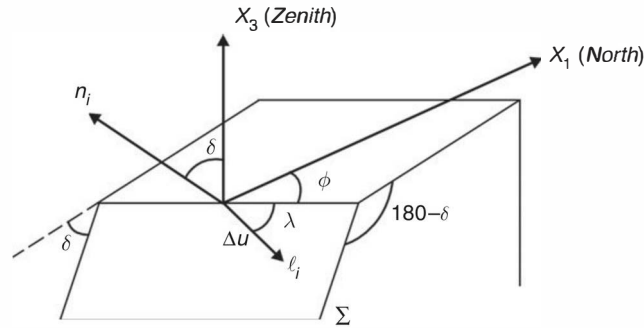


Fig. 17.2 Parameters of the motion on a fault.

faults for which, $\lambda = 0^\circ$, 180° or -180° , the motion is horizontal, for $\lambda = 90^\circ$ or -90° , the motion is vertical, and, for other values of λ , the motion has vertical and horizontal components, of reverse or normal type according to its sign. The *slip* or *displacement*, Δu , is the distance traveled in the relative motion of a point on one side of the fault with respect to one on the other. If Δu varies along the fault plane, its mean value is $\Delta \bar{u}$. The area of the fault is S (for a rectangular fault $S = LD$, where L is its length and D is its width; for a circular fault $S = \pi a^2$, where a is its radius). Thus, the orientation of the motion of the fault is given by the three angles ϕ , δ , and λ , and its dimensions are given by its area S and mean displacement $\Delta \bar{u}$.

In Reid's model of elastic rebound, faulting is caused by the sudden release of accumulated elastic strain on a fault when the strength of the material or the friction that keeps the fault locked are overcome (Fig. 17.1). Fracturing can be approached in two different ways, kinematic and dynamic. *Kinematic models* of the source consider the slip of the fault Δu without relating it to the stresses that cause it. Fracturing is described purely in terms of the slip vector $\Delta \mathbf{u}(x_i, t)$ as a function of the coordinates on the fault plane and time. From models of this type, it is relatively simple to determine the corresponding elastic displacement field. The second approach, *dynamic models*, considers the complete fracture process relating the fault slip to the stress acting on the focal region. A complete dynamic description must be able to describe fracturing from the properties of the material of the focal region and the stress conditions. Dynamic models present greater difficulties and their solution, in many cases, can be found only by numerical methods.

17.2 Equivalent forces. Point sources

The first mathematical formulation of the mechanism of earthquakes was presented by Hiroshi Nakano (1923) using the ideas already developed in elastodynamic theory by Love (1892, 1927) and Lamb (1904). Nakano used the point-source approximation and represented the source by a system of body forces acting at a point. Since these forces must represent the fracture phenomenon they are called *equivalent forces*.

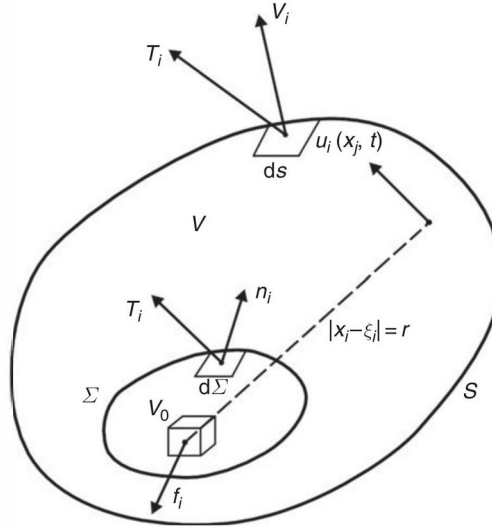


Fig. 17.3

The stress and displacement in a medium of volume V surrounded by a surface S . The body force $\mathbf{F}$ and stress $\mathbf{T}$ are acting in the focal region of volume V_0 , surrounded by the surface Σ .

The problem may be stated as follows. Let us consider an elastic medium of volume V surrounded by a surface S . In its interior there is a small region of volume V_0 , the focal region, surrounded by a surface Σ , where fracture takes place (Fig. 17.3). This process can be represented by a distribution of internal body forces $\mathbf{F}(\xi_i, t)$ acting per unit volume inside V_0 . Internal forces must satisfy the condition that their resulting sum is null. If it is assumed that no other body forces (gravity, etc.) are present, the equation of motion (4.62) can be written as

$$\int_{V-V_0} [\rho \ddot{u}_i(x_i, t) - \tau_{ij,j}(x_j, t)] dV = \int_{V_0} F_i(\xi_i, t) dV, \quad (17.1)$$

where ξ_i are the coordinates inside the focal region and x_i are those outside. Only elastic displacements and stresses outside the focal region are considered. From the static case in (4.65), the body forces F_i are formally related to stresses inside V_0 by

$$F_i = -\tau_{ij,j}. \quad (17.2)$$

In the case of a point source (Fig. 17.4), if the volume V is an infinite medium, the equation (17.1), according to (4.65), is given by

$$\rho \ddot{u}_i - \tau_{ij,j} = F_i, \quad (17.3)$$

where F_i are forces acting at a point that is selected as the origin of the x_i coordinates where the elastic displacements $\mathbf{u}(x_i, t)$ are evaluated. These forces are the limit of the forces acting on $V_\bullet$ as it tends to zero:

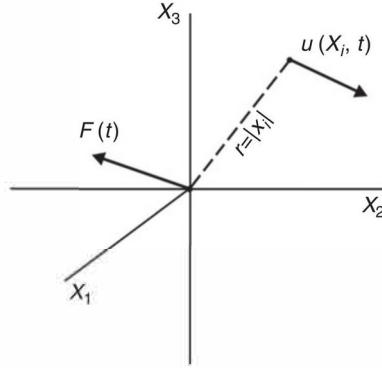


Fig. 17.4 A body point force $F(t)$ acting at the origin of coordinates and elastic displacements $u(x, t)$ at a point at a distance r .

$$F_i(t) = \lim_{V_0 \rightarrow 0} \int_{V_0} F_i(\xi_k, t) dV. \quad (17.4)$$

For a homogeneous medium, equation (17.3) can be expressed in terms of displacements, using (4.67), as

$$\rho \ddot{u}_i - C_{ijkl} u_{k,lj} = F_i. \quad (17.5)$$

This equation allows determination of the elastic displacements $u(x, t)$ produced by a force or system of forces F acting at the origin of the coordinates. In the inverse problem, from the observed elastic displacements we can obtain certain characteristics of these forces.

17.2.1 Formulation using Green's function

A more convenient formulation of the problem can be obtained by using the representation theorem in terms of Green's function (Section 4.8). According to (4.99), if the internal body forces F_k are limited to the focal region V_0 (Fig. 17.5) and on its surface Σ the stresses and displacements are null, we obtain for a volume V surrounded by a surface S that

$$u_i(x_s, t) = \int_{-\infty}^{\infty} d\tau \int_{V_0} F_k(\xi_p, \tau) G_{ki}(\xi_p, \tau; x_s, t) dV + \int_{-\infty}^{\infty} d\tau \int_S T_j(\eta_p) G_{ji}(\eta_p, \tau; x_s, t) - u_j(\eta_p) C_{jlmn} G_{mi,n}(\eta_p, \tau; x_s, t) v_l dV, \quad (17.6)$$

where x_s is the point where the displacements are evaluated at time t , ξ_p and η_p are generic points at the focal region V_0 and the external surface S , respectively, τ is the time dependence of the force, $T_i = \tau_{ij} v_j$ is the stress vector acting on S , v_i is the normal to the surface element dS , and G_{ki} is Green's function of the medium, defined by equation (4.94) in Section 4.8. Green's function, a tensor, is continuous throughout the volume V and represents the effect of propagation in the medium. As we saw in Section 4.8, Green's function is the solution of the equation of motion for an impulsive force and depends on the characteristics (C_{ijkl} , ρ , and boundary conditions) of the medium. If the medium is infinite,

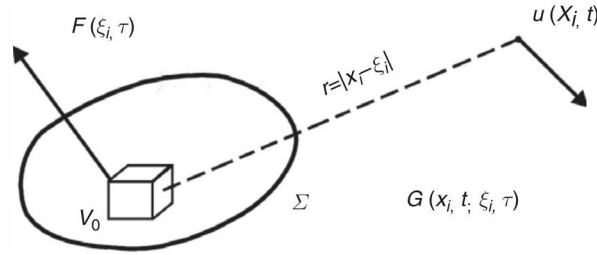


Fig. 17.5

A body force $F(\xi_i, t)$ acting on a source volume V_0 and elastic displacements $u(x, t)$ at a distance r in an infinite medium.

conditions on the surface S are homogeneous (the stress and displacement are null) and equation (17.6) becomes (Fig. 17.5)

$$u_i(x_s, t) = \int_{-\infty}^{\infty} d\tau \int_{V_0} F_k(\xi_s, \tau) G_{ki}(x_s, t; \xi_s, \tau) dV. \quad (17.7)$$

As explained in Section 4.8, the function G_{ki} acts as a “propagator” of the effects of forces F_k from the point where they are acting (ξ_s inside V_0) to point x_s outside V_0 where the elastic displacements u_i are produced. For a point focus at the origin of coordinates, we have

$$u_i(x_s, t) = \int_{-\infty}^{\infty} F_k(\tau) G_{ki}(x_s, t - \tau) d\tau. \quad (17.8)$$

The elastic displacements are now given by the time convolution of the forces acting at the focus with Green’s function of the medium.

From the point of view of representation of the seismic source in terms of equivalent forces, there are two ways to find the elastic displacements corresponding to a point source. The first consists of solving directly the equation of motion (17.5). This implies solving a second-order inhomogeneous differential equation for displacements or solving a homogeneous equation and introducing the forces as boundary conditions. In both cases, the problem is not easy. The second consists of using equation (17.8). In this case we need to have prior knowledge of Green’s function. Since Green’s function is the solution of the equation of motion, this equation must be solved anyway. However, the advantage of the second approach is that, for a given medium, the equation of motion must be solved only once to find Green’s function, whereas in the first, it must be solved for each system of forces. Thus, in the point-source problem, the first approach requires for each system of forces solution of equation (17.5) in the same medium. In the second approach, equation (4.94) is solved only once to find Green’s function. Then, for each system of forces, we apply equation (17.8), which is a convolution of each system of forces with Green’s function.

17.2.2 Single and double couples

Several systems of forces have been proposed to represent the source of an earthquake. For point sources, the most common are those of a couple of forces (SC, a single couple) and

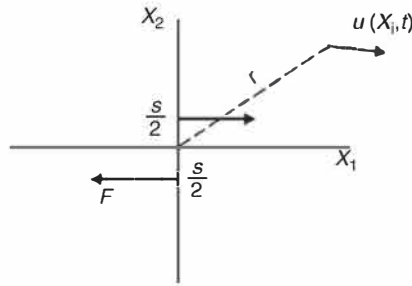


Fig. 17.6 A single couple of forces acting at the origin and elastic displacements $u(x, t)$ at a distance r .

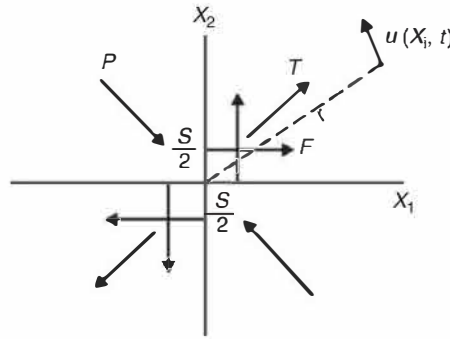


Fig. 17.7 A double couple of forces without a moment acting at the origin and the displacement $u(x, t)$ at a distance r . An equivalent system of two linear dipoles of pressure and tension forces P and T is shown.

two couples perpendicular to each other without a resulting moment (DC, a double couple) (Figs. 17.6 and 17.7). The second system is also equivalent to two linear dipoles of forces (with arms in the same direction as the forces) corresponding to pressure and tension acting at 45° to the couples. Both models were thought to represent shear fracture, but, as we will see, this is true only for the second. Besides, internal sources must have a zero resulting force and moment, which is not the case for the single couple. For models of extended sources, distributions of single or double couples on a plane surface were used.

For a point source, elastic displacements due to a couple of forces can be derived from those from a single force. If u_i^1 are the displacements due to a force acting at the origin in the x_1 direction, those of a couple of forces in the plane (x_1, x_2) with forces along the x_1 axis and the arm in the x_2 direction (Fig. 17.6) are derived by performing a Taylor expansion for each force of the couple, displaced by $s/2$ from the origin along the x_2 axis. For the force in the positive direction of x_1 and shifted by $s/2$ from the origin in the positive direction of x_2 , the elastic displacement is

$$u_i^+ = u_i^1 + \frac{s}{2} u_{i,2}^1. \quad (17.9)$$

For the force in the negative direction of x_1 and shifted by $s/2$ in the negative direction of x_2 , the displacement is

$$u_i^- = -u_i^1 + \frac{s}{2}u_{i,2}^1. \quad (17.10)$$

The displacement due to a single couple is the sum of the two:

$$u_i^{\text{SC}} = su_{i,2}^1. \quad (17.11)$$

For a double couple in the (x_1, x_2) plane, with forces in the directions of x_1 and x_2 (Fig. 17.7), using (17.11), the elastic displacement is given by

$$u_i^{\text{DC}} = s(u_{i,2}^1 + u_{i,1}^2). \quad (17.12)$$

If we substitute (17.11) into (17.8), we obtain the displacement due to a SC, as defined above, in terms of Green's function:

$$u_i^{\text{SC}} = \int_{-\infty}^{\infty} M(\tau) G_{i1,2}(t - \tau) d\tau, \quad (17.13)$$

where $M(t) = F(t)s$ is the moment of the couple. For a DC in the x_1 and x_2 directions, using (17.12), we obtain

$$u_i^{\text{DC}} = \int_{-\infty}^{\infty} M(\tau) [G_{i1,2}(t - \tau) + G_{i2,1}(t - \tau)] d\tau. \quad (17.14)$$

For a SC in an arbitrary orientation constant with time, with forces in the direction of the unit vector $\mathbf{l}$ and the arm in that of $\mathbf{n}$, where $\mathbf{n} \cdot \mathbf{l} = 0$, and a DC with the second couple with forces in the direction of $\mathbf{n}$ and its arm in that of $\mathbf{l}$, general expressions are

$$u_i^{\text{SC}} = \int_{-\infty}^{\infty} M l_k n_l G_{ik,l} d\tau, \quad (17.15)$$

$$u_i^{\text{DC}} = \int_{-\infty}^{\infty} M (l_k n_l + n_k l_l) G_{ik,l} d\tau. \quad (17.16)$$

Let us consider now two perpendicular linear dipoles with opposite signs. The linear dipole with forces in the positive direction corresponds to tension, whereas that with forces in the negative direction corresponds to pressure. If the forces are in the directions of x_1' and x_2' , in a similar form to those in (17.12), the elastic displacements are

$$u_i'^{\text{TP}} = s(u_{i,1}'^1 - u_{i,2}'^2). \quad (17.17)$$

If the system of coordinates (x_1', x_2') in (17.17) is rotated by 45° with respect to (x_1, x_2) of equation (17.12), the two expressions can be shown to be equivalent. If the tension and pressure forces are defined by the scalar moment M and unit vectors $\mathbf{T}$ and $\mathbf{P}$, in a similar form to those in (17.16), the elastic displacement in terms of Green's function is

$$u_i^{\text{TP}} = \int_{-\infty}^{\infty} M (T_k T_l - P_k P_l) G_{ik,l} d\tau. \quad (17.18)$$

For the two equivalent systems, relations between the unit vectors $\mathbf{P}$ and $\mathbf{T}$ and $\mathbf{n}$ and $\mathbf{l}$ are

$$\mathbf{P} = \frac{1}{\sqrt{2}}(\mathbf{n} - \mathbf{l}), \quad (17.19)$$

$$\mathbf{T} = \frac{1}{\sqrt{2}}(\mathbf{n} + \mathbf{l}), \quad (17.20)$$

$$\mathbf{B} = \mathbf{n} \times \mathbf{l} = \mathbf{P} \times \mathbf{T}, \quad (17.21)$$

where $\mathbf{B}$ is the unit vector normal to the plane of the forces. This vector is known as the *null axis*, since there is no component of forces in its direction. In this way, for a DC point source, we can define two orthogonal systems of axes in the directions of the unit vectors $\mathbf{n}$, $\mathbf{l}$, and $\mathbf{B}$ and $\mathbf{P}$, $\mathbf{T}$, and $\mathbf{B}$ to specify the orientation of the source. We will see that the second system corresponds to the principal axes of stress.

17.3 Fractures and dislocations

If an earthquake is produced by fracturing of the Earth's crust in faults, then we can use a mechanical representation of its source in terms of fractures or *dislocations* in an elastic medium. The theory of elastic dislocations was developed by Volterra in 1907 and discussed by Love (1927). Its first applications to the problem of seismic sources were by Anna Viktorovna Vvedenskaya (1956), Vladimir Keylis-Borok (1956), J. A. Steketee (1958), Leon Knopoff and Freeman Gilbert (1960) and Robert Burridge and Knopoff (1964).

A dislocation consists of an internal surface inside an elastic medium across which there is a discontinuity of displacement or strain. Here we will consider only displacement dislocations, that is, those for which there is a discontinuity of displacement but the stress is continuous. The problem will be formulated using the representation theorem in terms of Green's function (Section 4.9). The focal region consists of an internal surface Σ with two sides (positive and negative). This surface can be considered to be derived from the focal volume V_0 that is flattened to form a surface with both sides together without any volume. The coordinates on this surface are ξ_i and the normals at the points are $n_i(\xi_k)$. From one side to the other of this surface there is a discontinuity in displacement or slip (Fig. 17.8), so that

$$u_i^+(\xi_k, t) - u_i^-(\xi_k, t) = \Delta u_i(\xi_k, t), \quad (17.22)$$

where the plus and minus signs refer to the displacements at each side of the surface Σ . If there are no body forces ($F = 0$), the stresses are continuous through Σ (their integral is null) and the conditions on the external surface S are homogeneous (all integrals on S are null), then equation (4.107) results in

$$u_n(x_s, t) = \int_{-\infty}^{\infty} d\tau \int_{\Sigma} \Delta u_i(\xi_s, \tau) C_{ijkl} n_j(\xi_s) G_{nk,l}(\xi_s, \tau; x_s, t) dS, \quad (17.23)$$

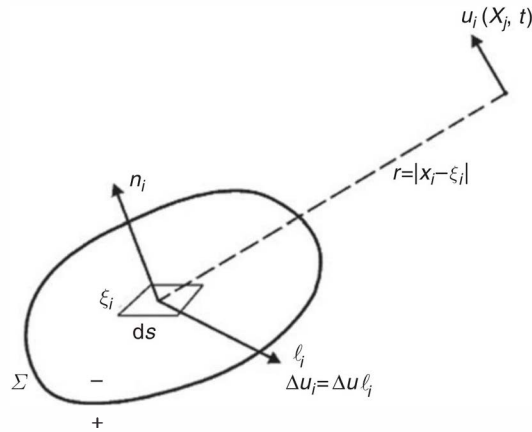


Fig. 17.8

A dislocation Δu on a surface Σ and the elastic displacement $\mathbf{u}(x, t)$ at a distance r in an infinite medium.

where x_j is the point where displacement is evaluated and t its time dependence, ξ_i a generic point on surface Σ and τ the time dependence of the slip Δu . In consequence, the seismic source is represented by a dislocation or discontinuity in displacement given by the slip vector Δu on the surface Σ , which corresponds to the relative displacement of the two sides of a fault. This is, then, an inelastic displacement that, once it has been produced, does not go back to the initial position. In the most general case, $\Delta u(\xi_i, \tau)$ can have a different direction for each point ξ_i of the surface Σ and, at each of these points, varies with time, starting from a zero value at $\tau = 0$, to a maximum value at a certain final time. The normal to the surface Σ , given by the unit vector $\mathbf{n}(\xi_i)$, can have different directions at points of the surface, but usually is considered to be constant; that is, Σ is a plane. Green's function $\mathbf{G}$ includes the effects of the medium on propagation from points ξ_i on the surface Σ to points x_j , where elastic displacements u_i are evaluated. To solve the problem, according to equation (17.23) we must first know the derivatives of Green's function for the medium, which are also known as *excitation functions*.

Equation (17.23) corresponds to a kinematic model of the source; that is, a model in which the elastic displacements $\mathbf{u}$ in the medium are derived from the slip vector Δu at the source, which represents the inelastic displacement of the two sides of a fault of surface Σ . The slip is assumed to be known rather than being derived from the stress conditions in the focal region as is done in dynamic models. In equation (17.23) the derivatives of Green's function include derivatives of the delta function. If we change the order of integration, integration over time of the product of Δu with the derivatives of the delta function results in time derivatives $\Delta \dot{u}$, that is, the *slip velocity*. Thus, elastic displacements depend not on the slip but rather on the slip velocity. This means that the source radiates elastic energy only while it is moving. When motion at the source stops it ceases to radiate energy ($\mathbf{u}$ is null).

As a particular case, let us consider an isotropic medium of elastic coefficients λ and μ , a plane surface Σ (with $\mathbf{n}$ constant), and a constant slip Δu with the same direction defined by the unit vector $\mathbf{l}$. The integrand of (17.23) becomes

$$\Delta u(t)[\lambda_k n_k \delta_{ij} + \mu(l_i n_j + l_j n_i)] G_{mij}. \quad (17.24)$$

The geometry of the source is now defined by the orientations of the two unit vectors $\mathbf{n}$ and $\mathbf{l}$. These two vectors, referring to the geographic system of axes (North, East, and nadir), define the orientation of the source, namely, $\mathbf{n}$ is the orientation of the fault plane and $\mathbf{l}$ is that of the slip (Fig. 17.2). Since they are unit vectors, each has only two independent components. For $i = j$, expression (17.24) gives the component of the displacement normal to the fault plane, which implies changes in volume, as it occurs, for example, in volcanic earthquakes. If $\mathbf{l}$ and $\mathbf{n}$ are perpendicular, the slip is along the fault plane, there are no changes in volume, and it represents a *shear fracture*, as it is the case of tectonic earthquakes. In this case ($\mathbf{n} \cdot \mathbf{l} = 0$), there are only three independent components of $\mathbf{n}$ and $\mathbf{l}$.

In the kinematic model of a dislocation on a plane surface with a constant slip, the parameters of the source are the elastic coefficients of the focal region λ and μ , four independent components of $\mathbf{n}$ and $\mathbf{l}$, defining the orientation of the fault plane and slip, the magnitude of the slip Δu , and the area S of the fault. There are, then, eight parameters that, when added to the four of the hypocenter (ϕ_0 , λ_0 , h , and t_0), sum up to 12. If the source is a shear fracture the number of parameters is only eleven.

17.4 The Green function for an infinite medium

The problem of determining Green's function is not an easy one and depends on the characteristics of each medium. As we saw in Section 4.8, Green's function is the solution of the equation of motion for an impulsive force in time and space (4.94) and (4.95). Let us consider Green's function for an infinite, homogeneous, isotropic elastic medium with force acting at the origin and $t = 0$. According to (4.95), Green's function is the solution of the equation

$$\rho \ddot{G}_{ij} - (\lambda + \mu) G_{ik,kj} - \mu G_{ij,kk} = \delta(x_s) \delta(t) \delta_{ij}. \quad (17.25)$$

In terms of the velocities α and β (4.71), we also have

$$\ddot{G}_{ij} - \alpha^2 \nabla (\nabla \cdot G_{ij}) + \beta^2 \nabla \times (\nabla \times G_{ij}) = \frac{1}{\rho} \delta(x_s) \delta(t) \delta_{ij}. \quad (17.26)$$

17.4.1 The radial force

We will start with a simplified problem in only one dimension for a radial force. In a problem with spherical symmetry and an impulsive radial force, elastic displacements have only radial components and the problem can be solved in one dimension (Fig. 17.9). Green's function has only one component, $G_{ij} = u(r, t)$. Since for this case $\nabla \times \mathbf{u} = 0$, equation (17.26) becomes

$$\ddot{u}(r, t) - \alpha^2 \nabla^2 u(r, t) = \frac{1}{\rho} \delta(r) \delta(t). \quad (17.27)$$

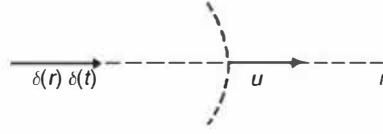


Fig. 17.9

An impulsive radial force at the origin and the displacement at a distance r .

On substituting the Laplacian in spherical coordinates for a dependence on r alone (A2.30), we have

$$\frac{\partial^2 u}{\partial t^2} = \frac{\alpha^2}{r} \frac{\partial^2}{\partial r^2} (ru) + \frac{1}{\rho} \delta(r) \delta(t). \quad (17.28)$$

If there are no forces, equation (17.28) for a harmonic dependence results in the Helmholtz equation (5.96), and its solution (5.99) is

$$u(r, t) = \frac{A}{r} \exp \left[i\omega \left(t - \frac{r}{\alpha} \right) \right]. \quad (17.29)$$

Let us now consider Poisson's equation (Morse and Feshbach, 1953),

$$\nabla^2 u(r) = -\frac{1}{\alpha^2 \rho} \delta(r). \quad (17.30)$$

Its solution is

$$u(r) = \frac{\delta(r)}{4\pi\alpha^2\rho r}. \quad (17.31)$$

We can consider that the solution of (17.28), in analogy with (17.29) and (17.31), can be written as

$$u(r, t) = \frac{\delta(t - r/\alpha)}{4\pi\alpha^2\rho r}. \quad (17.32)$$

Then, for an impulsive radial force, the elastic radial displacements have the form of an impulse that depends on $(t - r/\alpha)$. This impulse travels with a velocity α (P wave), arrives at a distance r at time $t = r/\alpha$, and its amplitude decreases with distance as $1/r$. Equation (17.32) represents Green's function for a radial force with spherical symmetry in an infinite homogeneous medium.

17.4.2 An impulsive force in an arbitrary direction

The complete problem of Green's function corresponds to an impulsive force in an arbitrary direction (Aki and Richards, 1980). Using Cartesian coordinates, we start with the particular case of a force in the x_1 direction applied at the origin of coordinates (Fig. 17.10). Equation (17.26), putting $G_{ij} = u_j$, is

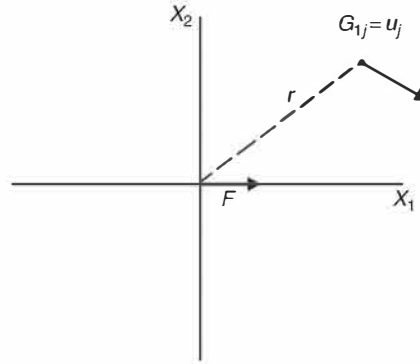


Fig. 17.10

An impulsive force F acting at the origin in the x_1 direction and the elastic displacement $\mathbf{u}$ equivalent to Green's function G_{1j} .

$$\ddot{\mathbf{u}} = F/\rho + \alpha^2 \nabla(\nabla \cdot \mathbf{u}) - \beta^2 \nabla \times (\nabla \times \mathbf{u}), \quad (17.33)$$

where F is given by

$$F = \delta(t)\delta(x_1, x_2, x_3)(1, 0, 0). \quad (17.34)$$

Just as in Section 4.7, we introduce the potentials ϕ and ψ for displacements and Φ and Ψ for forces ((4.85) and (4.88)),

$$\mathbf{u} = \nabla\phi + \nabla \times \psi, \quad (17.35)$$

$$\mathbf{F} = \nabla\Phi + \nabla \times \Psi. \quad (17.36)$$

On replacing (17.35) and (17.36) into (17.33) we obtain (4.89) and (4.90),

$$\ddot{\phi} - \alpha^2 \nabla^2 \phi = \Phi/\rho, \quad (17.37)$$

$$\ddot{\psi} - \beta^2 \nabla^2 \psi = \Psi/\rho. \quad (17.38)$$

We introduce now a new vectorial function W from which we can deduce the force potentials:

$$\Phi = -\nabla \cdot W, \quad (17.39)$$

$$\Psi = \nabla \times W. \quad (17.40)$$

By substitution into (17.36), we obtain

$$\nabla^2 W = -F. \quad (17.41)$$

This equation has the form of Poisson's equation (17.30) and its solution can be given as

$$W(\xi_i) = \frac{F}{4\pi q}, \quad (17.42)$$

where $q = |\xi_i|$ is the distance from the origin, where the force is applied to a generic point of coordinates ξ_i . If $\mathbf{F}$ is distributed in a volume V , wherein the coordinates of each point are ξ_i , the solution of (17.41) is

$$\mathbf{W}(\xi_i) = \frac{1}{4\pi} \int_V \frac{\mathbf{F}(\xi_i)}{q} dV. \quad (17.43)$$

In our case, $\mathbf{F}$ is given by (17.34) and $\mathbf{W}$ is

$$\mathbf{W} = \frac{\delta(t)}{4\pi q} (1, 0, 0). \quad (17.44)$$

According to (17.39), (17.40), and (17.44), the force potentials Φ and Ψ are

$$\Phi(\xi_i, t) = \frac{\delta(t)}{4\pi} \frac{\partial}{\partial \xi_1} \left(\frac{1}{q} \right), \quad (17.45)$$

$$\Psi(\xi_i, t) = \frac{\delta(t)}{4\pi} \left[0, \frac{\partial}{\partial \xi_3} \left(\frac{1}{q} \right), -\frac{\partial}{\partial \xi_2} \left(\frac{1}{q} \right) \right]. \quad (17.46)$$

Equations (17.37) and (17.38) for ϕ and ψ have the same form as (17.27); therefore, their solutions can be written in a similar form to (17.32), where the right-hand terms are Φ/ρ and Ψ/ρ , respectively. Since the force potentials are defined in the whole volume V , the solutions must be written in integral form as in (17.43):

$$\phi(x_i, t) = \frac{1}{4\pi\alpha^2\rho} \int_V \frac{\Phi(t - r'/\alpha)}{r'} dV, \quad (17.47)$$

$$\psi(x_i, t) = \frac{1}{4\pi\beta^2\rho} \int_V \frac{\Psi(t - r'/\beta)}{r'} dV, \quad (17.48)$$

where $r' = |x_i - \xi_i|$, with x_i where displacement u_i is evaluated. In our problem, the volume V represents the whole of the infinite space. We must notice that all potentials are defined over the whole volume under consideration. This may sound strange, especially for force potentials. Even when forces are acting at a point (the origin of coordinates), the potentials Φ and Ψ that represent them are defined over the whole volume of the problem, in our case the infinite space. By substituting expression (17.45) into (17.47), we obtain the solution of the displacement potential ϕ for an impulsive force in the x_1 direction:

$$\phi(x_i, t) = \frac{1}{16\pi^2\rho\alpha^2} \int_V \frac{\delta(t - r'/\alpha)}{r'} \frac{\partial}{\partial \xi_1} \left(\frac{1}{q} \right) dV, \quad (17.49)$$

where V represents the infinite medium. A similar expression is found for ψ by using (17.46) and (17.48). The evaluation of these integrals is performed in the following manner. The integral over the volume V (in our case the infinite medium) is separated into two parts, first an integral over a spherical surface S , with its center at x_i , the point where the displacements are evaluated, and a radius equal to $\alpha\tau$ for ϕ and to $\beta\tau$ for ψ (Fig. 17.11), and second an integral over the time τ from zero to infinity. In this form we cover the infinite medium. With these substitutions equation (17.49) becomes

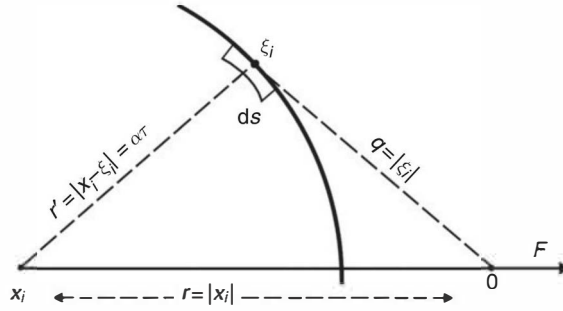


Fig. 17.11

A spherical surface with its center at x_i with radius $r' = |x_i - \xi_i| = \alpha\tau$ for the evaluation of integral (17.49) (ξ_i is the variable of integration and x_i where the displacement u_i is evaluated).

$$\phi(x_i, t) = \frac{1}{16\pi^2\rho\alpha^2} \int_0^\infty \frac{\delta(t-\tau)}{\tau} \int_S \frac{\partial}{\partial \xi_1} \frac{1}{q} dS d\tau, \quad (17.50)$$

where $q = |\xi_i|$ is the distance from the origin to an element dS of the spherical surface, S , $|x_i| = r$ is the distance from the origin to the point where $\mathbf{u}$ is evaluated, and $r' = |x_i - \xi_i| = \alpha\tau$ is the distance from the point x_i to a point ξ_i on the spherical surface (Fig. 17.11). We change the orders of differentiation and integration in the integral over the surface S . Then, it can be shown that, when the spherical surface includes the origin ($\tau > r/\alpha$), the integral over S of $1/q$ is constant and equal to 4π , and consequently its derivative is zero. If the surface does not include the origin ($\tau < r/\alpha$), the integral has the value $4\pi\alpha^2\tau^2/r$ (Aki and Richards, 1980). Thus, we obtain

$$\int_S \frac{\partial}{\partial \xi_1} \frac{1}{q} dS = 4\pi\alpha^2\tau^2 \frac{\partial}{\partial \xi_1} \frac{1}{r},$$

and, on substituting this into (17.50), we have

$$\phi = \frac{1}{4\pi\rho} \frac{\partial}{\partial x_1} \frac{1}{r} \int_0^{r/\alpha} \tau \delta(t-\tau) d\tau. \quad (17.51)$$

The derivative with respect to ξ_1 now becomes one with respect to x_1 and the integral over τ extends only to $\tau = r/\alpha$ since, for $\tau > r/\alpha$, the integrand is zero. In a similar manner we obtain the expression for the vector potential ψ , using (17.48) and (17.46):

$$\psi = \frac{1}{4\pi\rho} \left(0, \frac{\partial}{\partial x_3} \frac{1}{r}, -\frac{\partial}{\partial x_2} \frac{1}{r} \right) \int_0^{r/\alpha} \tau \delta(t-\tau) d\tau. \quad (17.52)$$

By taking the corresponding derivatives according to (17.35) in (17.51) and (17.52), we obtain the displacement $\mathbf{u}$, that is, Green's function for a force in the x_1 direction or G_{i1} . The derivatives of $1/r$ can be expressed in terms of the direction cosines γ_i of the vector $\mathbf{r}$ from a generic point ξ_i where the force is applied (in our case the origin) to the point x_i where Green's function is evaluated (Fig. 17.12). This is easily shown, since if

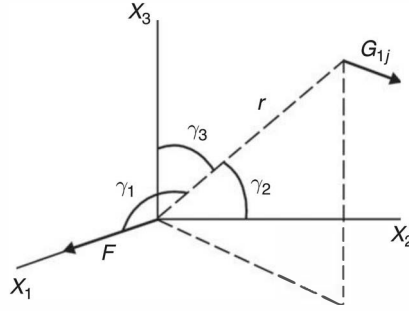


Fig. 17.12

The geometry of an impulsive force $\mathbf{F}$ in the x_1 direction and displacements (Green's function) G_{1j} at a distance r and direction cosines γ_i .

$$r = [(x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 + (x_3 - \xi_3)^2]^{1/2},$$

its derivative with respect to x_i is

$$\frac{\partial r}{\partial x_1} = \frac{x_1 - \xi_1}{r} = \cos(r, x_1) = \gamma_1,$$

and, in general,

$$\frac{\partial r}{\partial x_i} = \gamma_i. \quad (17.53)$$

The derivatives with respect to ξ_j are

$$\frac{\partial r}{\partial \xi_i} = -\gamma_i.$$

It is also easy to show that

$$\frac{\partial}{\partial x_j} \frac{1}{r} = -\frac{\gamma_j}{r^2}, \quad (17.54)$$

$$\frac{\partial \gamma_i}{\partial x_j} = -\frac{1}{r} (\gamma_i \gamma_j - \delta_{ij}). \quad (17.55)$$

After taking the corresponding derivatives and substituting in the direction cosines, we obtain

$$G_{i1} = \frac{1}{4\pi\rho} \left[\frac{1}{r^3} (3\gamma_1 \gamma_i - \delta_{1i}) \int_{r/\beta}^{r/\beta} \tau \delta(t - \tau) d\tau + \frac{1}{r\alpha^2} \gamma_1 \gamma_i \delta\left(t - \frac{r}{\alpha}\right) - \frac{1}{r\beta^2} (\gamma_1 \gamma_i - \delta_{1i}) \delta\left(t - \frac{r}{\beta}\right) \right]. \quad (17.56)$$

This equation corresponds to a force in the x_1 direction. We can generalize this result for Green's function corresponding to a force in an arbitrary direction (of unit vector v_j) denoted by the subindex j :

$$G_{ij} = \frac{1}{4\pi\rho} \left[\frac{1}{r^3} (3\gamma_i\gamma_j - \delta_{ij}) \int_{r/\alpha}^{r/\beta} \tau \delta(t - \tau) d\tau + \frac{1}{r\alpha^2} \gamma_i\gamma_j \delta\left(t - \frac{r}{\alpha}\right) - \frac{1}{r\beta^2} (\gamma_i\gamma_j - \delta_{ij}) \delta\left(t - \frac{r}{\beta}\right) \right]. \quad (17.57)$$

This is the expression for Green's function for an infinite, homogeneous, isotropic elastic medium with velocities α and β . This is a very important result in elastodynamics, which gives the elastic displacement field for the most fundamental type of source (an impulse in time and space). It constitutes the basic building block for seismic source studies.

A similar fundamental problem is the static solution for a constant force acting at a point in the direction of the unit vector v_j . For an infinite, homogeneous, isotropic elastic medium, the static displacements are solutions of the equation (4.97)

$$\alpha^2 \nabla(\nabla \cdot \mathbf{u}) - \beta^2 \nabla \times (\nabla \times \mathbf{u}) = \mathbf{F}/\rho.$$

The solution can be found in a similar way to that in the previous problem by expressing the displacement and force in terms of potentials. This leads to equations of the form of Poisson's equation (Udias *et al.*, 2014). The result for the displacement, in terms of the direction cosines γ_i , is

$$S_{ij} = \frac{F}{8\pi\rho r} \left[\left(\frac{1}{\beta^2} - \frac{1}{\alpha^2} \right) \gamma_i\gamma_j + \left(\frac{1}{\beta^2} + \frac{1}{\alpha^2} \right) \delta_{ij} \right]. \quad (17.58)$$

The subindex j indicates the direction of the force and, just like in (17.57), the displacements are given by a tensor. This expression, is Somigliana's tensor for an infinite, homogeneous, isotropic elastic medium, a fundamental equation in elastostatics.

17.5 The separation of near and far fields

The first term of Green's function in equation (17.57) depends on the distance as r^{-3} and the other two have r^{-1} dependence. Thus, the displacement represented by the first term is attenuated more rapidly with distance and for this reason is called the *nearfield*. This term depends on both α and β , and is a displacement of mixed P and S motion. The second and third terms constitute the *farfield*, where P and S waves are separated. If we are interested in the strong ground motion (accelerations) near the focus of an earthquake we use the near and far fields, while for larger distances we use only the far field. In both cases, near and far fields, displacements have two parts: one, called the *radiation pattern*, that depends on the direction cosines and expresses the spatial distribution of amplitudes; and another that depends on time or the *wave form*.

17.5.1 The near field

The time dependence of the near field (17.57) can be rewritten (Knopoff, 1967) as

$$\int_{r/\alpha}^{r/\beta} \tau \delta(t - \tau) d\tau = \int_{-\infty}^{\infty} \tau \delta(t - \tau) \left[H\left(\tau - \frac{r}{\alpha}\right) - H\left(\tau - \frac{r}{\beta}\right) \right] d\tau, \quad (17.59)$$

where $H(t)$ is the *step* or *Heaviside function* (named after Oliver Heaviside). According to the properties of step and delta functions, the integral on the right-hand side of (17.59) results in

$$\int_{-\infty}^{\infty} [\] d\tau = tH\left(t - \frac{r}{\alpha}\right) - tH\left(t - \frac{r}{\beta}\right). \quad (17.60)$$

Each of these two terms can be written as

$$tH\left(t - \frac{r}{\alpha}\right) = \left(t - \frac{r}{\alpha}\right)H\left(t - \frac{r}{\alpha}\right) + \frac{r}{\alpha}H\left(t - \frac{r}{\alpha}\right). \quad (17.61)$$

The term $(t - r/\alpha)H(t - r/\alpha)$ is a *ramp function* of slope unity. The complete expression for the near field can be written as

$$\begin{aligned} G_{ij}^{\text{NF}} = \frac{1}{4\pi\rho} (3\gamma_i\gamma_j - \delta_{ij}) & \left\{ \frac{1}{r^3} \left[\left(t - \frac{r}{\alpha}\right)H\left(t - \frac{r}{\alpha}\right) - \left(t - \frac{r}{\beta}\right)H\left(t - \frac{r}{\beta}\right) \right] \right. \\ & \left. + \frac{1}{r^2} \left[\frac{1}{\alpha}H\left(t - \frac{r}{\alpha}\right) - \frac{1}{\beta}H\left(t - \frac{r}{\beta}\right) \right] \right\}. \end{aligned} \quad (17.62)$$

The time dependence of the near field now has two parts that depend on distance as r^{-3} and r^{-2} , both depending on the velocities of P and S waves. The first part is the difference between two ramp functions and the second is the difference between two step functions of different amplitudes. The result is shown in Fig. 17.13. The part that depends on r^{-3} is formed by a ramp of unit slope starting at $t = r/\alpha$ and continuing until $t = r/\beta$. From this time onward the displacement has a constant amplitude of $1/\beta - 1/\alpha$. The part that depends on r^{-2} is a step function starting at $t = r/\alpha$ and amplitude $1/\alpha$, followed at $t = r/\beta$ by one of amplitude $1/\alpha - 1/\beta$. The displacement in the near field has a part that remains constant with time. The radiation pattern is common to the complete near-field displacement.

17.5.2 The far field

The far field (the part that depends on $1/r$) of Green's function is formed by separate P and S waves:

$$G_{ij}^P = \frac{1}{4\pi\rho\alpha^2 r} \gamma_i\gamma_j \delta\left(t - \frac{r}{\alpha}\right), \quad (17.63)$$

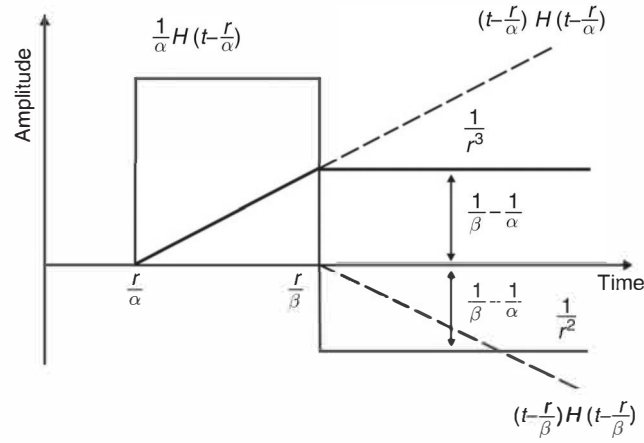


Fig. 17.13 The near-field displacement of Green's function for an infinite medium as a function of time at a distance r .

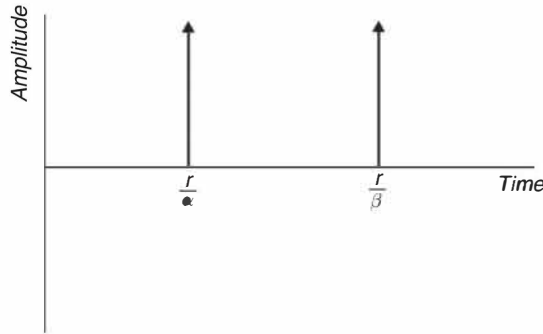


Fig. 17.14 The far-field displacement of Green's function for an infinite medium (P and S waves) at a distance r .

$$G_{ij}^S = \frac{-1}{4\pi\rho\beta^2 r} (\gamma_i \gamma_j - \delta_{ij}) \delta\left(t - \frac{r}{\beta}\right). \quad (17.64)$$

The time dependence is in both cases a delta function (Section 4.8). The far field is, then, formed by two impulses that propagate with velocities α and β , that is, P and S waves of impulsive form (Fig. 17.14).

The radiation patterns for P and S waves are different ((17.63) and (17.64)). To represent the radiation patterns we take polar coordinates (r, θ) with their center at the focus and consider the distribution of normalized amplitudes. If the force is in the x_1 direction, the normalized components of the displacement of P waves in the (x_1, x_3) plane are given (Fig. 17.15) by

$$G_{11}^P = \gamma_1 \gamma_1 = \cos \theta \cos \theta, \quad (17.65)$$

$$G_{31}^P = \gamma_3 \gamma_1 = \cos \theta \sin \theta. \quad (17.66)$$

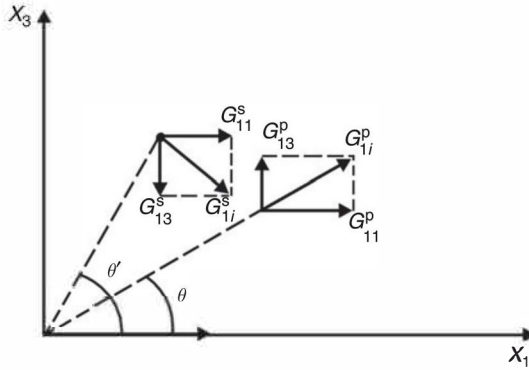


Fig. 17.15 Components of Green's function in the far field corresponding to P and S waves for a force in the x_i direction.

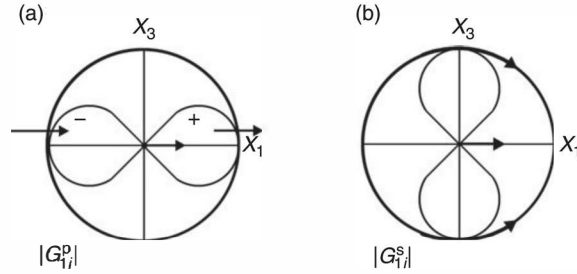


Fig. 17.16 The radiation pattern of Green's function on the (x_1, x_3) plane for a force in the x_1 direction: (a) P waves and (b) S waves.

It can easily be seen that the displacement is in the radial direction, as expected for P waves, and that its modulus is

$$|G_{ii}^p| = \cos \theta. \quad (17.67)$$

For S waves, the normalized displacement components are

$$G_{11}^s = -(\gamma_1 \gamma_1 - 1) = \sin \theta \sin \theta, \quad (17.68)$$

$$G_{31}^s = -\gamma_1 \gamma_3 = -\sin \theta \cos \theta. \quad (17.69)$$

The resulting displacement is in the transverse direction and its modulus is

$$|G_{ii}^s| = \sin \theta. \quad (17.70)$$

The displacements of P and S waves are in the radial and transverse directions, correspondingly, for each type of wave. By giving values from 0 to 360° to θ , we obtain radiation patterns. In both cases the pattern has two lobes (Fig. 17.16). For P waves, the displacements are radial, in the right lobe outward, that is, *compressions* and in the left inward, *dilations*, with maxima at $\theta = 0$ and π (Fig. 17.16a). For S waves, the displacements are transverse which converge toward the direction of the force in both lobes with maxima at

$\theta = \pi/2$ and $3\pi/2$ (Fig. 17.16b). P waves have a nodal plane (x_2, x_3) normal to the force and S waves are in the plane (x_1, x_2) that contains the force.

Green's function for an infinite homogeneous isotropic medium corresponds to the most simple type of propagating medium. This can be useful as a first approximation in some cases, but more complex models are necessary to represent the structure of the Earth. In the first place a free surface must be considered so that surface waves are generated. Better approximations are found using layered models of plane and spherical media. There are several methods for solving the problem, two of the most commonly used are the *reflectivity* and *normal mode* methods (Müller, 1985; Kennett, 2001, 2003). If we use complex Earth models this will result in more complex Green's functions.

17.6 A shear dislocation or fracture. The point source

Let us consider the seismic source represented by a shear dislocation fracture, with fault plane Σ of area S , normal n , and slip $\Delta u(\xi, \tau)$ in the direction of the unit vector l , contained on the plane so that l and n are perpendicular ($n \cdot l = 0$). For an infinite, homogeneous isotropic medium, according to (17.23) and (17.24), the expression for the displacements $u_k(x_i, t)$ is

$$u_k = \int_{-\infty}^{\infty} d\tau \int_{\Sigma} \Delta u \mu (l_i n_j + l_j n_i) G_{ki,j} dS. \quad (17.71)$$

If the distance from the observation point to the source is large in comparison with the source dimensions ($r \gg \Sigma$) and the wave lengths are also large ($\lambda \gg \Sigma$), the problem can be approximated by a point source, and equation (17.71) takes the form,

$$u_k(x_i, t) = \mu S (l_i n_j + l_j n_i) \int_{-\infty}^{\infty} \Delta u(\tau) G_{ki,j}(t - \tau) d\tau. \quad (17.72)$$

The displacements are given by the time convolution of the slip with the derivatives of Green's function.

For the far field, Green's functions for P and S waves are given by (17.63) and (17.64). The derivatives for P waves are

$$G_{ki,j}^p = \frac{1}{4\pi\rho\alpha^2} \frac{\partial}{\partial \xi_j} \left[\frac{1}{r} \gamma_i \gamma_k \delta\left(t - \frac{r}{\alpha}\right) \right]. \quad (17.73)$$

If in the derivatives we keep only the terms that depend on the least negative power of r ($1/r$), then we obtain

$$\frac{1}{r} \gamma_i \gamma_k \frac{\partial}{\partial \xi_j} \delta\left(t - \frac{r}{\alpha}\right) = -\frac{1}{r\alpha} \gamma_i \gamma_k \delta\left(t - \frac{r}{\alpha}\right) \frac{\partial r}{\partial \xi_j}.$$

On substituting in the direction cosine, we obtain for P waves,

$$G_{kij}^P = \frac{1}{4\pi\rho\alpha^3 r} \gamma_i \gamma_k \gamma_j \dot{\delta}\left(t - \frac{r}{\alpha}\right). \quad (17.74)$$

In a similar form, for S waves,

$$G_{kij}^S = \frac{-1}{4\pi\rho\beta^3 r} (\gamma_i \gamma_k - \delta_{ik}) \gamma_j \dot{\delta}\left(t - \frac{r}{\beta}\right). \quad (17.75)$$

This approximation is consistent with that of the far field. Now, we substitute (17.74) and (17.75) into (17.72) and take into account the property of the derivative of the delta function:

$$\int_{-\infty}^{\infty} \Delta u(\tau) \dot{\delta}\left(t - \frac{r}{\alpha} - \tau\right) d\tau = \Delta \dot{u}\left(t - \frac{r}{\alpha}\right). \quad (17.76)$$

The final result, after this substitution, gives for the displacements of the P and S waves in the far field,

$$u_j^P = \frac{\mu S}{4\pi\rho\alpha^3 r} (n_k l_i + n_i l_k) \gamma_i \gamma_k \gamma_j \Delta \dot{u}\left(t - \frac{r}{\alpha}\right), \quad (17.77)$$

$$u_j^S = \frac{\mu S}{4\pi\rho\beta^3 r} (n_k l_i + n_i l_k) (\delta_{ij} - \gamma_i \gamma_j) \gamma_k \Delta \dot{u}\left(t - \frac{r}{\beta}\right). \quad (17.78)$$

It is important to notice, as we mentioned in Section 17.3, that elastic displacements depend on the slip velocity or rate of slip $\Delta \dot{u}$. The source radiates elastic energy only while it is moving and ceases so to do when it stops. If the source time function is a step function, $\Delta u(t) = \Delta u H(t)$, its derivative is the delta function, and we obtain from (17.77) and (17.78),

$$u_j^P = \frac{M_0}{4\pi\rho\alpha^3 r} (n_k l_i + n_i l_k) \gamma_i \gamma_k \gamma_j \delta\left(t - \frac{r}{\alpha}\right), \quad (17.79)$$

$$u_j^S = \frac{M_0}{4\pi\rho\beta^3 r} (n_k l_i + n_i l_k) (\delta_{ij} - \gamma_i \gamma_j) \gamma_k \delta\left(t - \frac{r}{\beta}\right), \quad (17.80)$$

where we have substituted in the seismic moment $M_\bullet = \mu \Delta u S$ (16.33). Therefore, for a step source time function, elastic displacements in the far field for P and S waves are impulses that arrive at a distance r at times $t = r/\alpha$ and $t = r/\beta$.

17.6.1 The radiation pattern

The radiation pattern consists of the spatial distribution of amplitudes around the source. Let us consider a shear fracture on the plane (x_1, x_2) , with slip in the x_1 direction, that is, $\mathbf{n} = (0, 0, 1)$ and $\mathbf{l} = (1, 0, 0)$ (Fig. 17.17). In a similar form to that for Green's function, the normalized displacements in the (x_1, x_3) plane in polar coordinates (r, θ) for P waves are

$$u_1^P = 2\gamma_1^2\gamma_3 = \sin(2\theta)\cos\theta, \quad (17.81)$$

$$u_3^P = 2\gamma_3^2\gamma_1 = \sin(2\theta)\sin\theta, \quad (17.82)$$

and those for S waves are

$$u_1^S = (1 - 2\gamma_1^2)\gamma_3 = -\cos(2\theta)\sin\theta, \quad (17.83)$$

$$u_3^S = (1 - 2\gamma_3^2)\gamma_1 = \cos(2\theta)\cos\theta. \quad (17.84)$$

We can see that as expected the displacement of P waves is in the radial direction and that of S waves is in the transverse direction. If we define the components u_r and u_θ in these two directions, we obtain

$$u_r^P = \sin(2\theta), \quad (17.85)$$

$$u_\theta^S = \cos(2\theta). \quad (17.86)$$

In both cases, the radiation pattern has four lobes or quadrants. For P waves the lobes have alternating directions of motion, outward or positive (compression) and inward or negative (dilation). There are two nodal planes, (x_1, x_2) and (x_3, x_2) , the first corresponds to the *fault plane* and the second, corresponding to the normal to this and to the direction of $\Delta\mathbf{u}$, is called the *auxiliary plane*. The maxima of the displacement are at 45° to the directions of $\mathbf{n}$ and $\mathbf{l}$ (Fig. 17.18a). In the four lobes of the radiation pattern of S waves, the motion changes direction. The maxima coincide with the directions of $\mathbf{n}$ and $\mathbf{l}$ and the nodal planes are at 45° to them (Fig. 17.18b). In both cases, the radiation pattern is symmetric and we can interchange $\mathbf{n}$ and $\mathbf{l}$ without changing the result. This is a consequence of the expression (17.71) being symmetric with respect to $\mathbf{n}$ and $\mathbf{l}$. For this reason, the radiation patterns of P and S waves do not distinguish the fault plane from the auxiliary plane. This ambiguity is present in the point source methods used to determine the orientation of the fault plane from far-field displacements of P and S waves.

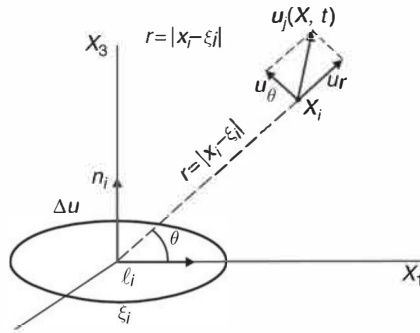


Fig. 17.17

Displacements $\mathbf{u}(\mathbf{x}, t)$ at a distance r due to a shear dislocation with slip in the x_1 direction on a plane normal to the x_3 axis.

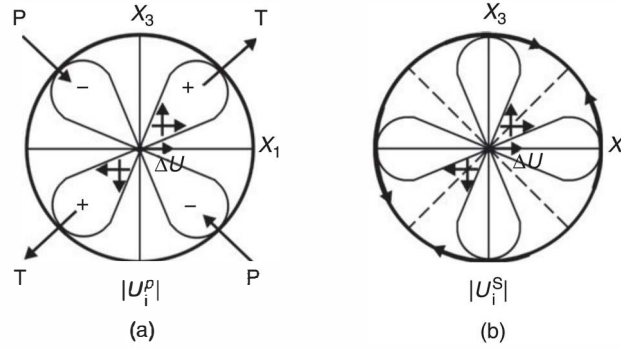


Fig. 17.18

The radiation pattern in the plane (x_1, x_3) due to a point shear dislocation (Fig. 17.17) and the equivalent DC and PT system of forces: (a) P waves and (b) S waves.

To study the radiation pattern in three dimensions we use spherical coordinates (r, θ, ϕ) . The focus is located at the center of a sphere of unit radius and displacements are evaluated for points on its surface. A system of Cartesian coordinates (x_1, x_2, x_3) with its origin at the center of the sphere, called the *source system*, is defined so that (x_1, x_2) is the fault plane, $\mathbf{n}$ is in the x_3 direction, and $\mathbf{l}$ is in that of x_1 . For a point on the surface of the sphere, the direction cosines of r with respect to the three axes are

$$\gamma_1 = \sin \theta \cos \phi, \quad (17.87)$$

$$\gamma_2 = \sin \theta \sin \phi, \quad (17.88)$$

$$\gamma_3 = \cos \theta, \quad (17.89)$$

where θ is measured from x_3 and ϕ is measured from x_1 . At each point of the spherical surface, we define a system of Cartesian coordinates with unit vectors (e_r, e_θ, e_ϕ) in the directions of the increments of r , θ , and ϕ . The components of displacements in these directions correspond to P waves and two components of S waves, respectively (Fig. 17.19). The normalized amplitudes are given by

$$P : \quad u_r = \sin(2\theta) \cos \phi, \quad (17.90)$$

$$S1 : \quad u_\theta = \cos(2\theta) \cos \phi, \quad (17.91)$$

$$S2 : \quad u_\phi = \cos \theta \sin \phi. \quad (17.92)$$

Displacements of P waves have two nodal planes, $\theta = \pi/2$, (x_1, x_2) , and $\phi = \pi/2$, (x_2, x_3) . Displacements of the S1 component have a nodal plane for $\phi = \pi/2$, (x_2, x_3) , and are null also for points of intersection of the surface of the focal sphere and the solid angle $\theta = \pi/4$ and $3\pi/4$. Displacements of S2 have two nodal planes for $\theta = \pi/2$, (x_1, x_2) , and $\phi = 0$, (x_1, x_3) .

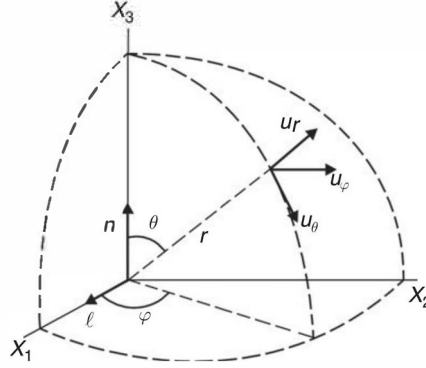


Fig. 17.19

Components of the displacement $\mathbf{u}(r, \theta, \phi)$ in spherical coordinates for the point shear dislocation of Fig. 17.17.

17.6.2 The geometry of a shear fracture

As we have seen, the orientation of a shear fracture is given by two orthogonal unit vectors $\mathbf{n}$ and $\mathbf{l}$. These two vectors must be expressed in relation to the geographic reference system defined by the axes x_1 , x_2 , and x_3 , positive in North, East, and nadir directions. With respect to these axes we define now the spherical coordinates θ measured from x_3 and ϕ measured from x_1 on the (x_1, x_2) (horizontal) plane. In reference to this system the vector $\mathbf{n}$ is given by

$$n_1 = \sin \theta_n \cos \phi_n, \quad (17.93)$$

$$n_2 = \sin \theta_n \sin \phi_n, \quad (17.94)$$

$$n_3 = \cos \theta_n. \quad (17.95)$$

We do this in a similar way for the vector $\mathbf{l}$ (l_1, l_2, l_3). Since $\mathbf{l}$ and $\mathbf{n}$ are orthogonal unit vectors, only three components are independent. We saw in Section 17.1 that a fault or shear fracture can also be defined by the angles ϕ , δ , and λ (Fig. 17.2). These angles can be expressed in terms of those defining the geographic orientations of the vectors $\mathbf{l}$ and $\mathbf{n}$ (Fig. 17.20):

$$\phi = -\phi_n + \pi/2, \quad (17.96)$$

$$\delta = \theta_n, \quad (17.97)$$

$$\lambda = \sin^{-1} \left(\frac{\cos \theta_l}{\sin \theta_n} \right). \quad (17.98)$$

The components of $\mathbf{n}$ and $\mathbf{l}$ referring to the geographic axes (x_1 , x_2 , and x_3) can be written in terms of ϕ , δ , and λ , in the form,

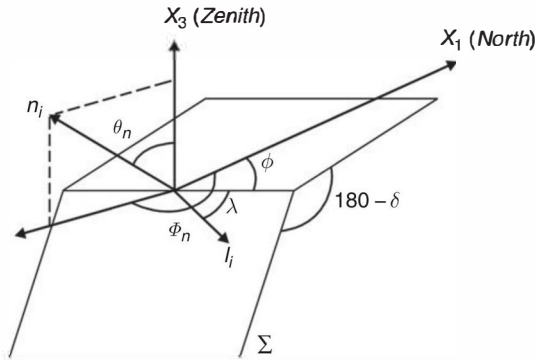


Fig. 17.20

Relations between the angles ϕ , δ , and λ , and θ_n and ϕ_n for a shear fault.

$$n_1 = -\sin \delta \sin \phi, \quad (17.99)$$

$$n_2 = \sin \delta \cos \phi, \quad (17.100)$$

$$n_3 = -\cos \delta, \quad (17.101)$$

$$l_1 = \cos \lambda \cos \phi + \cos \delta \sin \lambda \sin \phi, \quad (17.102)$$

$$l_2 = \cos \lambda \sin \phi - \cos \delta \sin \lambda \cos \phi, \quad (17.103)$$

$$l_3 = -\sin \lambda \sin \delta. \quad (17.104)$$

In every case, the orientation of the source is given uniquely by three parameters, namely, ϕ , δ , and λ , or θ_n , ϕ_n , and θ_l . Since, as we have mentioned already, there is always an ambiguity with respect to the vectors $\mathbf{n}$ and $\mathbf{l}$, the source orientation given by (θ_n, ϕ_n) and (θ_l, ϕ_l) does not assume a distinction between fault and auxiliary planes. If we know from other information, for example field geology or aftershock distribution, the orientation of the fault plane, this can be given in terms of ϕ , δ , and λ . If we do not know which one is the fault plane, we must give values of ϕ , δ , and λ for both nodal planes.

Displacement of P waves and of SV and SH components of S waves can be referred to the geographic coordinate axes through the direction of the seismic ray. In a sphere surrounding the focus, this is a straight line from the center to the surface. If ϕ' is the azimuth of the ray measured from North and i is the take-off angle of the ray measured from the downward vertical, then the direction cosines of the ray with respect to the geographic axes (North, East, and nadir) are

$$\gamma_1 = \sin i \cos \phi', \quad (17.105)$$

$$\gamma_2 = \sin i \sin \phi', \quad (17.106)$$

$$\gamma_3 = \cos i. \quad (17.107)$$

The components of P, SV, and SH displacements along the geographic axes (Fig. 17.21) are

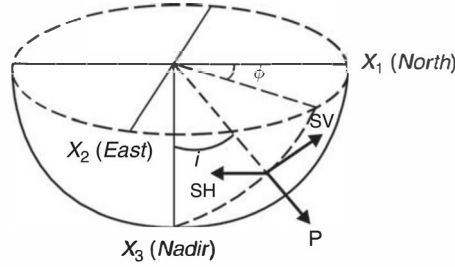


Fig. 17.21

Displacements of P waves and SV and SH components of S waves, in spherical geometry for a source at the center.

$$\begin{aligned}
 P_1 &= P \sin i \cos \phi', & SV_1 &= SV \cos i \cos \phi', & SH_1 &= -SH \sin \phi' \\
 P_2 &= P \sin i \sin \phi', & SV_2 &= SV \cos i \sin \phi', & SH_2 &= SH \cos \phi', \\
 P_3 &= P \cos i, & SV_3 &= -SV \sin i, & SH_3 &= 0
 \end{aligned}$$

where the P, SV, and SH wave amplitudes depend on the location of the observation point with respect to the source orientation. For points on the sphere, the observation point is given by (ϕ', i) and the orientation of the source is given by (ϕ, δ, λ) or $(\theta_n, \phi_n, \theta_1)$.

17.7 The source time function

The *source time function* (STF) $\Delta u(t)$ represents the slip's dependence on time and is an important characteristic of the focal mechanism. We have already considered the most simple STF, namely, the step function. There are other functions; some commonly used ones, including the step function (Fig. 17.22), are

$$\Delta u(t) = \Delta u H(t), \quad (17.108)$$

$$\Delta u(t) = \begin{cases} \Delta u \frac{t}{\tau}, & 0 \leq t \leq \tau, \\ \Delta u, & t > \tau \end{cases}, \quad (17.109)$$

$$\Delta u(t) = \Delta u H(t) (1 - e^{-t/\tau}). \quad (17.110)$$

In all cases, the slip of the fault starts at $t = 0$, and, once it has reached its maximum value Δu , it stays constant. After breaking, the fault does not return to its initial state. In the first case (17.108), $\Delta u(t)$ has the form of a step or Heaviside function such that the slip reaches its maximum value instantaneously at time $t = 0$. In the second case (17.109), $\Delta u(t)$ increases linearly from $t = 0$ to $t = \tau$, and at that time reaches its maximum value and stops moving. This STF introduces a new parameter of the source, namely, τ , the time taken for the slip to reach its maximum value, called the *rise time*. In the third case (17.110), $\Delta u(t)$ is a continuous function for $t > 0$. The slip reaches its maximum value asymptotically with time. For the rise time, $\Delta u(\tau) = 0.63 \Delta u$.

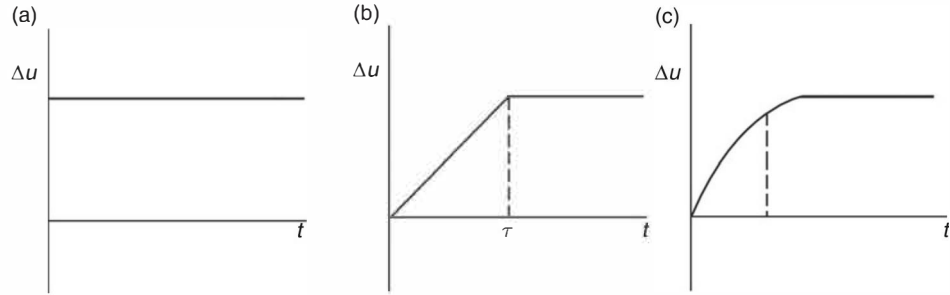


Fig. 17.22 The source time function for a slip $\Delta u(t)$: (a) a step function; (b) a ramp function with a rise time τ ; and (c) an exponential function with a rise time τ .

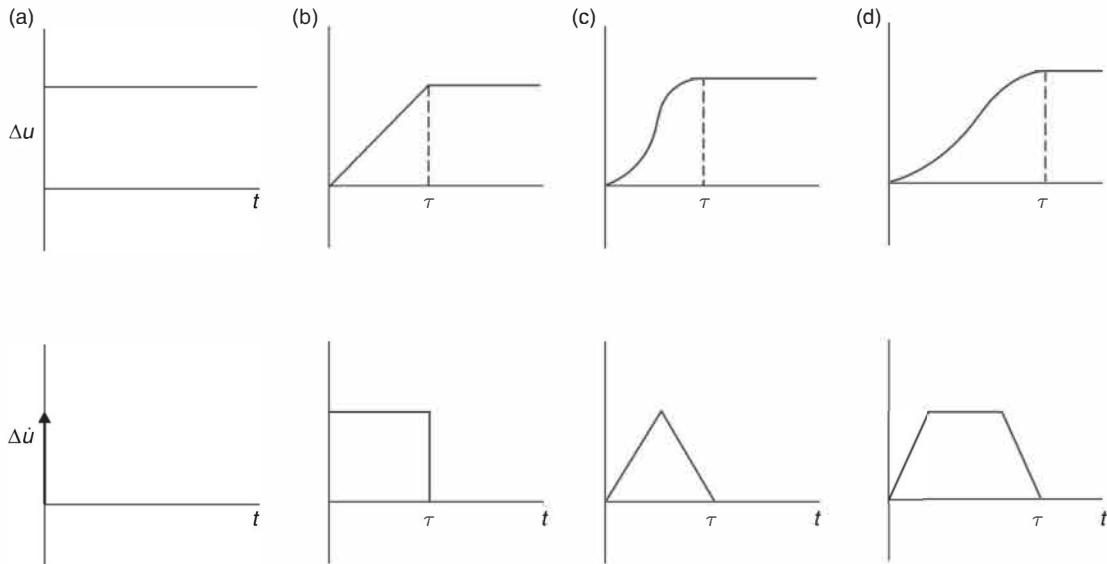


Fig. 17.23 The source time function for a slip velocity $\Delta \dot{u}(t)$ and its relation to that for a slip $\Delta u(t)$: (a) an impulsive function; (b) a rectangular function with duration τ ; (c) a triangular function of duration τ ; and (d) a trapezoidal function of duration τ .

We have seen in equations (17.77) and (17.78) that elastic displacements depend on the slip velocity $\Delta \dot{u}$. For this reason, the time dependence of the slip velocity or slip rate is often also called the STF. For the first two models ((17.108) and (17.109)), we obtain

$$\Delta \dot{u}(t) = \Delta v_M \delta(t), \quad (17.111)$$

$$\Delta \dot{u}(t) = \Delta v_M [H(t) - H(t - \tau)]. \quad (17.112)$$

In these two models, at $t = 0$, the slip velocity jumps instantaneously from 0 to its maximum value Δv_M (Figs. 17.23a and b). In the first, the slip velocity is an impulse and

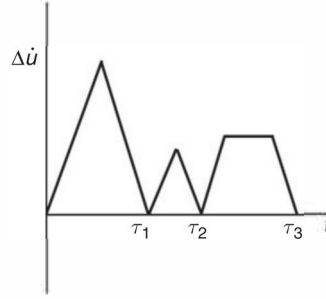


Fig. 17.24

The source time function for a slip velocity $\Delta \dot{u}(t)$ for a complex source with several (three) events.

in the second it has a duration τ with constant value. More realistic is to define a STF with a slip velocity that increases from zero to its maximum value and then decreases to zero after the rise time τ . A model that satisfies these conditions is a triangular function (Fig. 17.23c):

$$\Delta \dot{u} = \begin{cases} 0, & t < 0 \\ \Delta v_M \frac{2t}{\tau}, & 0 \leq t \leq \frac{\tau}{2} \\ \Delta v_M \frac{2(\tau - t)}{\tau}, & \frac{\tau}{2} < t \leq \tau \\ 0, & t > \tau \end{cases} \quad (17.113)$$

The slip velocity increases linearly from zero at $t = 0$ to reach its maximum value Δv_M at $t = \tau/2$ and then decreases to zero for $t = \tau$. During the first part of the process, the slip acceleration ($\Delta \ddot{u}$) is positive, whereas in the second it is negative. If we want to increase the duration of the source process we can use a STF of trapezoidal form (Fig. 17.23d). In this case, the slip velocity maintains its maximum value for a certain time before decreasing to zero at $t = \tau$.

The models of the STF we have mentioned represent simple sources consisting of a single event. A complex source can be represented by a STF consisting of several triangles or trapezoids of different heights (Fig. 17.24). In this way we can represent with a point source a mechanism that has several accelerations ($\Delta \ddot{u} > 0$), decelerations ($\Delta \ddot{u} < 0$), and stops ($\Delta \dot{u} = 0$), during the total process of fracturing. Each event has its rise-time and the *total fracture time* T is the sum of all of them (in Fig. 17.24, $T = \tau_1 + \tau_2 + \tau_3$). Note that in this case the total fracture time is different from the rise time.

17.8 The equivalence between forces and dislocations

We have seen that the source of earthquakes can be represented by systems of forces (17.7) or by displacement dislocations (17.23). We will see now that the double couple system of forces is equivalent to a shear dislocation or fracture. For a point source, the elastic

displacement due to a single couple with forces in the x_1 direction and the couple's arm s in the x_3 direction ($\mathbf{l} = (1, 0, 0)$ and $\mathbf{n} = (0, 0, 1)$), according to (17.15), is

$$u_i^{\text{SC}} = \int_{-\infty}^{\infty} M G_{i1,3} d\tau, \quad (17.114)$$

where $M = Fs$ is the moment of the couple. For a double couple with forces in the x_1 and x_3 directions ($\mathbf{l}_1 = (1, 0, 0)$, $\mathbf{n}_1 = (0, 0, 1)$, $\mathbf{l}_2 = (0, 0, 1)$, and $\mathbf{n}_2 = (1, 0, 0)$), by substitution into (17.16), we obtain

$$u_j^{\text{DC}} = \int_{-\infty}^{\infty} M [G_{i1,3} + G_{i3,1}] d\tau. \quad (17.115)$$

We substitute Green's function for the far-field P waves (17.63) into (17.114) and (17.115), and, putting $M(t) = MH(t)$, we obtain

$$u_i^{\text{DC:P}} = 2u_i^{\text{SC:P}} = \frac{M}{4\pi\rho\alpha^3 r} \gamma_1 \gamma_3 \gamma_r \delta\left(t - \frac{r}{\alpha}\right). \quad (17.116)$$

For displacements of S waves in the far field, by substituting (16.64) into (16.114) and (16.115), we get

$$u_i^{\text{SC:S}} = \frac{-M}{4\pi\rho\beta^3 r} (\gamma_i \gamma_1 \gamma_3 - \delta_{i1} \gamma_3) \delta\left(t - \frac{r}{\beta}\right), \quad (17.117)$$

$$u_i^{\text{DC:S}} = \frac{-M}{4\pi\rho\beta^3 r} (2\gamma_i \gamma_1 \gamma_3 - \delta_{i1} \gamma_3 - \delta_{i3} \gamma_1) \delta\left(t - \frac{r}{\beta}\right). \quad (17.118)$$

The radiation pattern for P waves is the same for SC and DC models. The first studies of focal mechanisms were based on the polarity distribution of P waves, which is the same for both models. This explains the discussion about which model really corresponded to the seismic source. The radiation pattern of S waves, however, is different for each model. Analysis of S waves, especially after 1960, proved that the DC model was the one that is consistent with observations.

Let us now consider elastic displacements in the far field due to a point shear dislocation (17.72). For the slip orientation given by $\mathbf{n} = (0, 0, 1)$ and $\mathbf{l} = (1, 0, 0)$, that is, a fracture on the plane (x_1, x_2) , with slip in the x_1 direction, equation (17.72) becomes

$$u_k = \mu S \int_{-\infty}^{\infty} \Delta u (G_{k1,3} + G_{k3,1}) d\tau. \quad (17.119)$$

This expression is equivalent to (17.115). In consequence, a double couple with one couple in the direction of the slip and the other perpendicular to the fault plane is equivalent to a shear fracture (Fig. 17.25). The same result is obtained when we compare expressions (17.79) and (17.80) for the far field of P and S waves due to a shear dislocation, substituting $\mathbf{l} = (1, 0, 0)$ and $\mathbf{n} = (0, 0, 1)$ with the expressions for a DC (17.116) and (17.118). In the comparison of the two sets of expressions, we find that the moment of the couple of forces

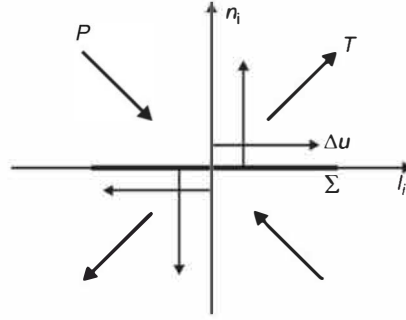


Fig. 17.25

The equivalence between a double couple (and **PT** system) and a point shear dislocation.

$M = Fs$ is equivalent to the seismic moment $M_0 = \mu \Delta u S$ of the shear fracture; in both cases the dimensions are force times distance (Nm).

Another more general way to show this equivalence is to start with the equality (Aki and Richards, 1980)

$$\int_V F_k G_{ik} dV = \int_\Sigma \Delta u \mu (n_k l_j + n_j l_k) G_{ikj} dS. \quad (17.120)$$

In this equation, we search for the set of body forces F_k distributed inside the volume V which corresponds to the shear dislocation on the surface Σ located inside V . Inside the volume V the coordinates are ξ_i and η_i on the surface Σ . We consider the same particular case for $\mathbf{n} = (0, 0, 1)$ and $\mathbf{l} = (1, 0, 0)$:

$$\int_V F_k G_{ik} dV = \int_\Sigma \mu \Delta u (G_{i1,3} + G_{i3,1}) dS. \quad (17.121)$$

The integral on the left-hand side can be expressed in the form of an integral over the volume V by using the delta function $\delta(\xi_i - \eta_i)$. This function is zero for all points ξ_i in V that do not coincide with the points η_i on the surface Σ , that is, for all points of the volume that are not on the surface. The left-hand side of equation (17.121) becomes

$$\int_V \mu \Delta u (G_{i1,3} + G_{i3,1}) \delta(\xi_n - \eta_n) dV. \quad (17.122)$$

According to the properties of the derivatives of the delta function, the derivatives of G_{ij} can be written using those of $\delta(\xi_i - \eta_i)$ in the form,

$$G_{ij,k} = \int_V \frac{\partial}{\partial \xi_k} [\delta(\xi_n - \eta_n)] G_{ij} dV. \quad (17.123)$$

On substituting (17.123) into (17.122), we obtain

$$\int_V \mu \Delta u \left(\frac{\partial \delta}{\partial \xi_3} G_{i1} + \frac{\partial \delta}{\partial \xi_1} G_{i3} \right) \delta(\xi_n - \eta_n) dV = \int_V F_k G_{ik} dV, \quad (17.124)$$

where both integrals are defined over the same volume. If we shrink the volume to a point, for a point source, the part in brackets can be taken out of the integral and the integral becomes

$$\int_V \mu \Delta u \delta(\xi_n - \eta_n) dV = \mu \Delta u S = M_0.$$

Then, equation (17.124) becomes

$$F_k G_{ik} = M_0 \left(\frac{\partial \delta}{\partial \xi_3} G_{i1} + \frac{\partial \delta}{\partial \xi_1} G_{i3} \right). \quad (17.125)$$

In this equation we can identify the components of the force F_k that are equivalent to the shear dislocation:

$$F_1 = M_0 \frac{\partial \delta}{\partial \xi_3}, \quad F_2 = 0, \quad F_3 = M_0 \frac{\partial \delta}{\partial \xi_1}. \quad (17.126)$$

Since the delta function represents an impulsive force, its derivatives with respect to a particular coordinate represent a dipole or a couple with its arm in that direction. For example, $\partial \delta / \partial x_3$ is a couple with its arm in the x_3 direction. Thus F_k in (17.126) is formed by two couples in the plane (x_1, x_3) , in the directions of x_1 and x_3 , or a double couple. The two couples have the same moment, equal to the seismic moment M_0 . In consequence, a double couple on the plane (x_1, x_3) is the force system equivalent to a shear fracture on the plane (x_1, x_2) , with slip in the x_1 direction. In general, the system of equivalent forces for a shear dislocation is a double couple with one couple in the direction of the slip and the other normal to the fault plane. Here we have shown this equivalence for a particular case and a point source. The equivalence can be shown in a more general form and also for an extended fracture. In the latter case, the equivalent forces are a distribution of double couples on the fault plane (Aki and Richards, 1980). Owing to this equivalence, the term double couple source is often used as a synonym for a shear fracture, which, as we have already seen, is the common mechanism for tectonic earthquakes.

17.9 Summary

As has already been advanced in Chapter 2, earthquakes are caused by breaking of fractures in the Earth's crust under tectonic stress. In this chapter we have presented the basic ideas about the earthquake mechanism, that is, mechanical models from which we can obtain an elastic displacement field, so that it can be compared with the observations. The simplest models of earthquake mechanisms are those in terms of point sources of equivalent forces. Systems of couples of forces, simple and double, and dipoles are used to represent faulting. These models can be extended to distributions of forces in volumes and on surfaces. A better representation of fracture is by dislocations in elastic media. The elastic displacement field produced by these sources is obtained using the representation theorem in terms of Green's function.

Green's function for an infinite homogeneous isotropic medium has been developed. Its solution gives the displacements produced by an impulsive force acting at a point in an arbitrary direction. The dependence of the displacements on distance can be separated into a near and a far field. For the far field the radiation patterns of P and S waves have been presented. Green's functions are used to find the elastic displacement corresponding to different kinds of sources.

Using Green's function for an infinite homogeneous medium, we obtain the elastic displacements owing to a point source shear dislocation or double couple. Radiation patterns for P and S waves and some different types of source time functions have been given. Finally we have shown the equivalence between a shear dislocation and a double couple source.

17.10 Problems

- P17.1 At the focus of an earthquake situated at depth 100 km, a point source consisting of an impulsive force of magnitude 10^{15} N is acting in the direction of the vertical axis. For a point on the surface 100 km distant and at azimuth 45° , find the components of the displacement of the P wave. Use displacements in far field with $\alpha = 6 \text{ km s}^{-1}$ and $\rho = 3.3 \text{ g cm}^{-3}$ and consider an infinite medium.
- P17.2 In an infinite medium there is a force of magnitude F with a harmonic time dependence acting in the negative direction of the vertical axis (x_3). Find the SH and SV components of the displacement for a point on the plane (x_1, x_3) in terms of the coordinate (r, θ), where θ is measured from x_3 .
- P17.3 An earthquake is caused by a shear fracture. The vectors n and l (normal and direction of slip) are, in terms of the angles Φ and Θ ,

$$\begin{aligned} n &= (57.13^\circ, 66.44^\circ) \\ l &= (305.96^\circ, 50.09^\circ). \end{aligned}$$

Calculate the orientation of the fault plane, the auxiliary plane, and the tension (T) and pressure (P) stress axes.

- P17.4 The displacement vector l of an earthquake is (0, 1, 0) and the vector normal to the plane of displacement n is (0, 1, 0). Determine:
- The components of the P wave displacement at the point of azimuth 45° and angle of incidence 30° .
 - The kind of mechanism it represents.
- P17.5 The focal mechanism of an earthquake can be represented by a double couple (DC) model. The orientation of the fault plane is azimuth 30° , dip 90° , and slip angle 0° .
- Calculate:
 - What kind of fault it is. Sketch it, indicating the direction of motion.
 - The auxiliary plane.
 - The azimuth of the stress axis.

- (b) A wave incident at a station has azimuth 180° and angle of incidence at the focus of 90° . Calculate the amplitude of the components of the P wave at that station.
- P17.6 Find the far-field amplitude displacements of P waves and S waves at a point at a distance $r = 1000$ km, with $i_h = 45^\circ$ and $a_z = 45^\circ$, for a shear-fracture point source with $\mathbf{n} = (0, 1, 0)$ and $\mathbf{l} = (1, 0, 0)$. The medium has $\rho = 3 \text{ g cm}^{-3}$, $\alpha = 6 \text{ km s}^{-1}$, and $\sigma = \frac{1}{4}$.

In [Chapter 16](#) we introduced the scalar seismic moment as a measure of the size of an earthquake. In formulation of the theory of the source mechanism, an important concept is that of the *seismic moment tensor*. The seismic moment tensor was first proposed by Gilbert (1970), who related it to the total drop in stress which is produced in the occurrence of earthquakes. Backus and M. Mulcahy (1976a, b) clarified that the moment tensor represents only that part of the internal drop in stress that is dissipated in inelastic deformations at the source. In the [previous chapter](#) we have presented the shear fracture and its equivalent force representation, the double couple, as models for the mechanism of earthquakes. The moment tensor allows a more general representation of the source, not limited to that model, for example, models in which there are changes in volume. Thus the moment tensor can be considered as a very general way of representing the mechanism of earthquakes. In this chapter we present the general theory of the seismic moment tensor restricted to the point source or the *first-order moment tensor*.

18.1 Definition of the moment tensor

The seismic moment tensor M_{ij} and the moment tensor density per unit volume or unit surface m_{ij} are defined as (Jost and Herrmann, 1989)

$$M_{ij} = \int_V m_{ij} dV = \int_S m_{ij} dS. \quad (18.1)$$

If we consider an elastic medium in which only elastic processes occur, then, in the absence of body forces, the equation of motion (4.65) is

$$\rho \ddot{u}_i = \tau_{ij,j}. \quad (18.2)$$

Since in the real or physical situation there are, besides elastic, also inelastic processes occurring, the total stress is given by σ_{ij} and τ_{ij} correspond only to the purely elastic model. The equation of motion for the real situation is, then,

$$\rho \ddot{u}_i = \sigma_{ij,j}. \quad (18.3)$$

If we define the moment tensor density m_{ij} as the stress in excess of the purely elastic stress or the *stress glut*, it will be given by the difference,

$$m_{ij} = \tau_{ij} - \sigma_{ij}. \quad (18.4)$$

If we substitute σ_{ij} from (18.4) into (18.3), we obtain

$$\rho \ddot{u}_j = \tau_{ij,j} - m_{ij,j}. \quad (18.5)$$

If we compare this equation with (18.3), in which the seismic source was represented by equivalent body forces F_i , we get

$$F_i = -m_{ij,j}. \quad (18.6)$$

Equivalent body forces can be derived from the moment tensor and both can be used to represent the seismic source. Equation (18.6) clarifies the meaning of (17.2), in which body forces were related to stress in the focal region. Equivalent body forces correspond only to the stress responsible for inelastic processes in the source region. The moment tensor according to (18.5) represents, precisely, those stresses that are directly related to inelastic displacements (fracture) at the source of an earthquake. Just like for equivalent body forces, the seismic moment tensor is defined only inside the focal region, the region where the inelastic processes take place, and is zero outside.

If we substitute m_{ij} , according to (18.6), into (17.7), we can express elastic displacements u_i outside the focal region in terms of m_{ij} and the corresponding Green function:

$$u_i = \int_{-\infty}^{\infty} d\tau \int_{V_0} -m_{kj,j} G_{ik} dV. \quad (18.7)$$

Integrating by parts with respect to the spatial coordinates gives

$$u_i = \int_{-\infty}^{\infty} m_{kj} G_{kj} d\tau + \int_{-\infty}^{\infty} d\tau \int_{V_0} m_{kj} G_{ik,j} dV.$$

In the absence of external forces and torques, the sum of all internal forces and moments is null; then, by an appropriate choice of the origin of coordinates, $m_{kj} G_{kj} = 0$, and we obtain

$$u_j = \int_{-\infty}^{\infty} d\tau \int_{V_0} m_{kj} G_{ik,j} dV. \quad (18.8)$$

If the moment tensor is defined only on a surface Σ , we use m_{kj} , the moment tensor density per unit surface, and we write (18.8) as a surface integral:

$$u_i = \int_{-\infty}^{\infty} d\tau \int_{\Sigma} m_{kj} G_{ik,j} dS. \quad (18.9)$$

Equations (18.8) and (18.9) show that elastic displacements outside the focal region can be derived from the seismic moment tensor and the derivatives of Green's function integrated over the focal region (V_0 or Σ). Since we have not specified its form, m_{ij} can represent a very general type of source. It corresponds to any system of internal body forces according to (18.6), provided that the nett effect of their sum and the sum of their moments is null. The moment tensor is, thus, a very convenient form in which to represent the source of an earthquake in a general way.

For a point source, equations (18.8) and (18.9) can be written in a compact form using an asterisk to express time convolution (Appendix 4):

$$u_i = M_{kj} * G_{ik,j}. \quad (18.10)$$

The physical meaning of the moment tensor can be understood in relation to the equivalent body forces. According to (17.7), the elastic displacements are given by

$$u_j(x_s, t) = \int_{-\infty}^{\infty} d\tau \int_{V_0} F_k(\zeta_s, \tau) G_{ik}(x_s, t; \zeta_s, \tau) dV. \quad (18.11)$$

If we perform a Taylor expansion of G_{ik} around the origin, $\zeta_k = 0$, the first three terms are

$$G_{ik}(\zeta_s) = G_{ik}(0) + \zeta_s \frac{\partial G_{ik}}{\partial \zeta_s} + \frac{1}{2} \zeta_n \zeta_s \frac{\partial^2 G_{ik}}{\partial \zeta_n \partial \zeta_s} + \dots \quad (18.12)$$

Taking only the first two terms and substituting into (18.11), the first term is zero by virtue of the condition that the sum of internal forces must be zero:

$$\int_{V_0} F_k(0) G_{ik}(0, x) dV = 0. \quad (18.13)$$

Therefore, we obtain

$$u_i = \int_{-\infty}^{\infty} d\tau \int_{V_0} F_k \zeta_j G_{ik,j} dV. \quad (18.14)$$

By comparison of (18.14) with (18.8), we find that

$$m_{jk} = \zeta_j F_k. \quad (18.15)$$

The moment tensor density m_{ij} corresponds to the first nonzero term in the Taylor expansion of (18.12), and thus it is called the *first-order moment tensor*. This term is associated with the first derivatives of Green functions. We can also derive moment tensors of higher order that are associated with higher derivatives of Green functions, for example, the second-order moment tensor, the third term in (18.12), which represents its variation with space. Higher-order moment tensors are related to the time duration and space characteristics of the source (Backus and Mulcahy, 1976a, b).

According to (18.15), the components of m_{ij} correspond to force couples or dipoles. The components m_{11} , m_{22} , and m_{33} are linear dipoles without moments, that is, the arm is in the same direction as the forces. The other components have their arms perpendicular to the forces and are couples with moments (Fig. 18.1). The condition for an internal source of zero net moment implies that the tensor is symmetric, $m_{ij} = m_{ji}$; couples with opposite moment must be equal. We have seen that Green's functions represent displacements due to impulsive forces, whereas their derivatives represent displacements due to couples or dipoles of impulsive forces. In consequence, according to equations (18.8) and (18.9), elastic displacements are given by the convolution of distributions of dipoles or couples of forces representing the source (moment tensor) with displacements due to couples of impulsive forces (derivatives of the Green's function).

The components of the moment tensor are expressed in relation to a coordinate system of reference, usually the geographic system, with its origin at the focus of the earthquake: for

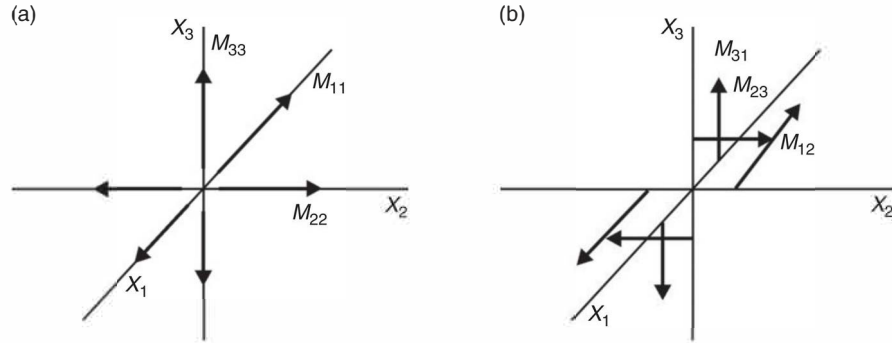


Fig. 18.1

Representations of the components of the seismic moment tensor: (a) for $i = j$ and (b) for $i \neq j$ (only three components are shown).

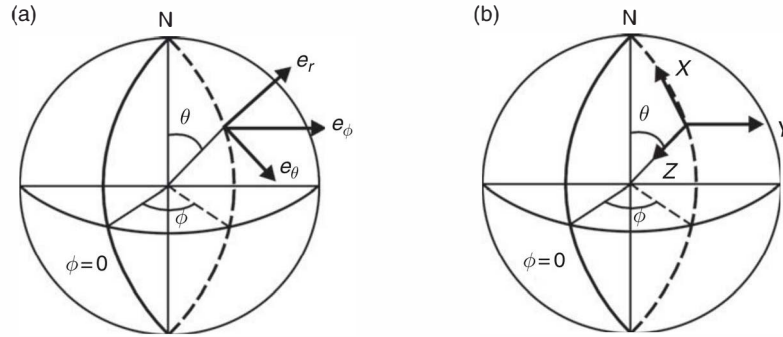


Fig. 18.2

Systems of coordinates at the focus referred to the geocentric geographic system of the Earth (r, θ, ϕ). (a) The system formed by unit vectors (e_r, e_θ, e_ϕ) (zenith, South, East). (b) The system in geographic directions (x, y, z) or (x_1, x_2, x_3) (North, East, nadir).

example, the Cartesian coordinate system (x_1, x_2, x_3) or (x, y, z), positive in the directions North, East, and nadir (Fig. 18.2b) or also North, West, and zenith. Another system that is also used is referred to as geocentric spherical coordinates of the focus (r, θ, ϕ), where r is in the radial direction, θ is the geocentric colatitude, and ϕ is the geocentric longitude. At the focus a Cartesian coordinate system with unit vectors e_r, e_θ , and e_ϕ (in the directions of positive increments of r, θ , and ϕ) is formed. This system has positive axes in the directions zenith, South, and East (Fig. 18.2a). The correspondence among the six components of the moment tensor in the three systems is

$$M_{11} = M_{xx} = M_{\theta\theta}$$

$$M_{22} = M_{yy} = M_{\phi\phi}$$

$$M_{33} = M_{zz} = M_{rr}$$

$$M_{12} = M_{xy} = -M_{\theta\phi}$$

$$M_{13} = M_{xz} = M_{r\theta}$$

$$M_{23} = M_{yz} = -M_{r\phi}.$$

18.2 The moment tensor and elastic dislocations

If we compare equations (18.9) and (17.23), we can define the moment tensor density corresponding to a dislocation with slip Δu on a surface Σ of normal $\mathbf{n}$ as

$$m_{ij} = C_{ijkl} \Delta u_k n_l, \quad (18.16)$$

and that for an isotropic medium as (17.24)

$$m_{ij} = \lambda n_k \Delta u_k \delta_{ij} + \mu (\Delta u_i n_j + \Delta u_j n_i). \quad (18.17)$$

If the slip direction is given by the unit vector $\mathbf{l}$, equation (18.17) becomes

$$m_{ij} = \Delta u [\lambda l_k n_k \delta_{ij} + \mu (l_j n_i + l_i n_j)]. \quad (18.18)$$

From this expression we can find the moment tensor for various types of sources, by specifying the orientations of $\mathbf{n}_i$ and $\mathbf{l}_i$.

18.2.1 An explosive source

An explosive source may be considered as an expansion along the three coordinate axes and may represent the source of a volcanic earthquake. This situation is represented by linear dipoles ($\mathbf{l}$ and $\mathbf{n}$ in the same direction) along each axis, that is, $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$. The moment tensor is the sum of the three and, using (18.18), we obtain (Fig. 18.3a)

$$m_{ij} = K \Delta u \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (18.19)$$

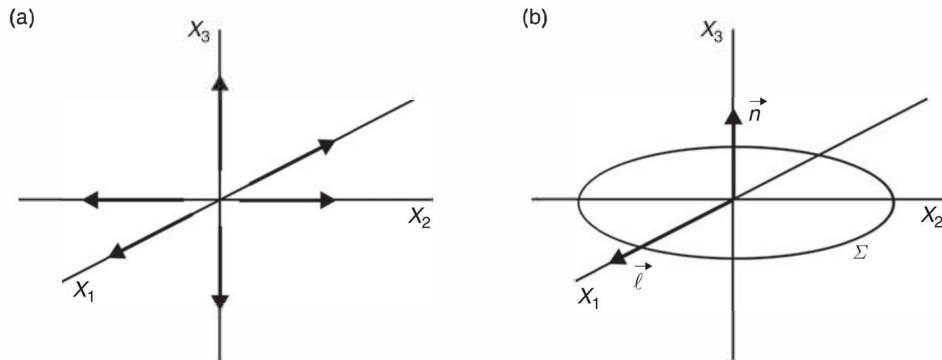


Fig. 18.3

Representations of the source: (a) for an explosion and (b) for a shear fracture.

where $K = \lambda + \frac{2}{3}\mu$ is the bulk modulus (4.24). The sum of elements of the principal diagonal gives the increase in volume per unit volume:

$$m_{11} + m_{22} + m_{33} = 3K \Delta u. \quad (18.20)$$

18.2.2 Shear fracture

In a shear fracture, slip $\Delta \mathbf{u}$ is along the fault plane; that is, $\mathbf{n}$ and $\mathbf{l}$ are perpendicular. Using equation (18.18) and the definition of M_0 , scalar seismic moment (Section 16.6), after integration over the source surface of area S , the moment tensor for a point source is

$$M_{ij} = M_0(l_i n_j + l_j n_i). \quad (18.21)$$

For a particular case in which the fault plane is the (x_1, x_2) plane, that is, $\mathbf{n} = (0, 0, 1)$, and the slip is in the x_1 direction, $\mathbf{l} = (1, 0, 0)$ (Fig. 18.3b), we obtain

$$M_{ij} = M_0 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \quad (18.22)$$

The sum of the principal diagonal is null, indicating that there is no change in volume.

The unit vectors $\mathbf{n}$ and $\mathbf{l}$ referring to the geographic reference system can be written in terms of the angles θ_n and ϕ_n , and θ_l and ϕ_l , respectively (Fig. 18.4). By substitution into (18.21) of the expressions for the components of $\mathbf{n}$, according to (17.93)–(17.95), and similarly for $\mathbf{l}$, we obtain for the six components of the normalized moment tensor,

$$\begin{aligned} m_{11} &= 2 \sin \theta_n \cos \phi_n \sin \theta_l \cos \phi_l \\ m_{22} &= 2 \sin \theta_n \sin \phi_n \sin \theta_l \sin \phi_l \\ m_{33} &= 2 \cos \theta_n \cos \theta_l \\ m_{12} &= \sin \theta_l \cos \phi_l \sin \theta_n \sin \phi_n + \sin \theta_l \sin \phi_l \sin \theta_n \cos \phi_n \\ m_{13} &= \sin \theta_l \cos \phi_l \cos \theta_n + \cos \theta_l \sin \theta_n \cos \phi_n \\ m_{23} &= \sin \theta_l \sin \phi_l \cos \theta_n + \sin \phi_n \sin \theta_n \cos \theta_l. \end{aligned} \quad (18.23)$$

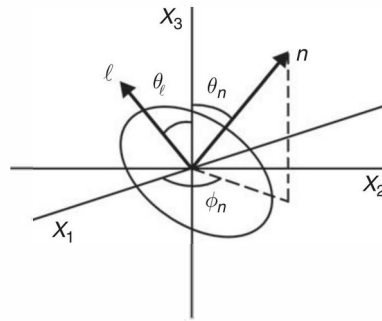


Fig. 18.4

Orientation of the unit vectors $\mathbf{l}$ and $\mathbf{n}$, and definitions of the angles θ_n , ϕ_n , and θ_l .

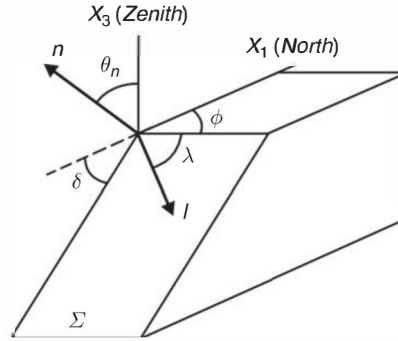


Fig. 18.5 Relation between unit vectors $\mathbf{l}$ and $\mathbf{n}$, and the angles ϕ , δ , and λ that define the motion in a fault.

As we have seen, a shear fracture can also be specified in terms of the angles ϕ , δ , and λ (Section 17.6). Using the relations between ϕ , δ , and λ , and n_i and l_i (Fig. 18.5) equations (17.99)–(17.104), from (18.21), the components of the moment tensor are

$$\begin{aligned}
 m_{11} &= -\sin \delta \cos \lambda \sin(2\phi) - \sin(2\delta) \sin^2 \phi \sin \lambda \\
 m_{22} &= \sin \delta \cos \lambda \sin(2\phi) - \sin(2\delta) \cos^2 \phi \sin \lambda \\
 m_{33} &= \sin(2\delta) \sin \lambda \\
 m_{12} &= \sin \delta \cos \lambda \cos(2\phi) + \frac{1}{2} \sin(2\delta) \sin(2\phi) \sin \lambda \\
 m_{13} &= -\sin \lambda \sin \phi \cos(2\delta) - \cos \delta \cos \lambda \cos \phi \\
 m_{23} &= \cos \phi \sin \lambda \cos(2\delta) - \cos \delta \cos \lambda \sin \phi.
 \end{aligned} \tag{18.24}$$

In (18.23), the expressions are symmetric with respect to $\mathbf{n}$ and $\mathbf{l}$, and imply that one cannot identify the fault plane from the two possible planes. Once we have selected one plane as the fault plane, the equations (18.24) relate the components of the moment tensor to the orientation of motion on the selected fault plane (Fig. 18.5). Naturally, the result is the same for values of ϕ , δ , and λ of the second plane.

18.3 Eigenvalues and eigenvectors

As we saw in the discussion of stress and strain tensors (Section 4.1), we can also perform an eigenvalue and eigenvector analysis of the moment tensor. Since this tensor is symmetric, its eigenvalues are real and its eigenvectors are mutually orthogonal. They satisfy the equation

$$(M_{ij} - \delta_{ij}\sigma)v_k = 0, \tag{18.25}$$

where the three eigenvalues σ_1 , σ_2 , and σ_3 are the roots of the cubic equation resulting from putting the determinant of (18.25) equal to zero. By substituting each eigenvalue into

(18.25) we obtain the three eigenvectors v_k^1 , v_k^2 , and v_k^3 , which form a set of three orthogonal axes, the *principal axes*. In reference to these axes, the moment tensor has the form

$$M_{ij} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}. \quad (18.26)$$

In this system, the moment tensor is formed by three linear dipoles in the direction of the principal axes and thus represents the *principal stresses*. If we order the eigenvalues $\sigma_1 > \sigma_2 > \sigma_3$, then σ_1 corresponds to the greatest stress, σ_3 to the least stress, and σ_2 to the intermediate stress. The sum of the elements of the principal diagonal is the first invariant of the tensor and has the same value for any reference system:

$$M_{11} + M_{22} + M_{33} = \sigma_1 + \sigma_2 + \sigma_3. \quad (18.27)$$

This sum represents the change in volume, as we saw for the explosive source. Thus we can define the *isotropic part* of the moment tensor as

$$\sigma_\bullet = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3). \quad (18.28)$$

We can also define the *differential stress*,

$$\sigma_\Delta = \sigma_1 - \sigma_3, \quad (18.29)$$

and the stress ratio,

$$R = \frac{\sigma_2 - \sigma_3}{\sigma_1 - \sigma_3}. \quad (18.30)$$

If we subtract the isotropic part from the tensor M_{ij} , we obtain the *deviatoric tensor* M'_{ij} , the sum of whose diagonal elements is always zero and does not include changes in volume:

$$M'_{ij} = M_{ij} - \delta_{ij}\sigma_\bullet. \quad (18.31)$$

According to (18.31), the moment tensor can be separated into two tensors, one isotropic, $M_{ij}^0 = \delta_{ij}\sigma_\bullet$, and the other deviatoric, M'_{ij} ,

$$M_{ij} = M_{ij}^0 + M'_{ij}. \quad (18.32)$$

Changes in volume are thus separated from other parts of the moment tensor. The moment tensor that represents an explosive source (18.19) is purely isotropic and that for a shear fracture (18.22) is purely deviatoric.

If we represent the moment tensor for an explosive source referring to its principal axis, we obtain the same result as that in (18.19). An explosive source is purely isotropic and any reference system is equivalent to the principal axes. For the shear fracture of (18.22) (the fault plane is normal to x_3 and slip is in the x_1 direction), the eigenvalues of the matrix (18.22) are 1, -1 , and 0. The eigenvectors are found by substitution into (18.25), resulting in $(1/\sqrt{2}, 0, 1/\sqrt{2})$ for $\sigma = 1$, $(1/\sqrt{2}, 0, -1/\sqrt{2})$ for $\sigma = -1$ and $(0, 1, 0)$ for $\sigma = 0$. The tensor referring to its principal axes is

$$M_{ij} = M_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (18.33)$$

Note that M_0 (scalar seismic moment) is the modulus of the moment tensor when it corresponds to a shear fracture (DC). This tensor represents two linear dipoles of positive and negative forces or tension and pressure forces along two of the principal axes, that is, in the (x_1, x_3) plane at 45° to the direction of slip (x_1). Thus, for this shear fracture, eigenvectors corresponding to the eigenvalues σ_1 and σ_3 define the principal axes of stress or pressure and tension axes, $\mathbf{P}$ and $\mathbf{T}$. The third axis corresponding to the zero eigenvalue is the null axis $\mathbf{B}$. In terms of the unit vectors $\mathbf{P}$ and $\mathbf{T}$, the moment tensor is given by

$$M_{ij} = M_0(T_i T_j - P_i P_j). \quad (18.34)$$

This result is analogous to that found in [Section 17.2.2](#) regarding the equivalence of a double couple to pressure and tension forces at a 45° angle to the couples.

18.4 Types of sources and separation of the moment tensor

We have already said that the moment tensor represents a very general type of source. The analysis of its eigenvalues indicates, in each case, the type of source. The most general case corresponds to three different eigenvalues, $\sigma_1 \neq \sigma_2 \neq \sigma_3$, whose sum is not zero, $\sigma_1 + \sigma_2 + \sigma_3 \neq 0$. Then, the source has changes in volume and, after separation of the isotropic part ([18.31](#)), the deviatoric part is of a general type, not necessarily a shear fracture or double couple.

If $\sigma_1 = \sigma_2 = \sigma_3$, as we have seen, the source is an isotropic expansion or contraction depending on the sign. In each case $\sigma_1 + \sigma_2 + \sigma_3$ represents the increase or decrease in volume. For positive sign the source represents an explosion.

For sources without nett volume changes, $\sigma_1 + \sigma_2 + \sigma_3 = 0$, the moment tensor is purely deviatoric. This condition is often imposed on sources of tectonic earthquakes. In this case, only two of the eigenvalues are independent, since $\sigma_2 = -\sigma_1 - \sigma_3$. For a shear fracture or double couple source, the moment tensor is deviatoric and must satisfy the conditions $\sigma_3 = -\sigma_1$ and $\sigma_2 = 0$.

Tectonic earthquake sources are commonly thought to be shear fractures or nearly so. However, this need not always be the case and even the possibility of the occurrence of changes in volume cannot be completely ruled out. Methods of inversion of the moment tensor from observations (see [Section 18.7](#)) do not impose any condition and result, for some earthquake sources, in the presence of certain amounts of volume change and non-double-couple components. For this reason, it is convenient to separate the moment tensor into three parts, one isotropic, corresponding to changes in volume, one of pure shear fracture or a double couple (DC), and a third that may be of various kinds (Strelitz, 1989). This analysis is called *partition* or *separation* of the moment tensor and can be expressed by writing

$$\mathbf{M} = \mathbf{M}^0 + \mathbf{M}^{DC} + \mathbf{M}^R. \quad (18.35)$$

The isotropic part M^0 (18.28) has already been defined. Partition of the deviatoric part ($M^{DC} + M^R$) can be done in several ways. The simplest is to separate this part into two DCs, *major* and *minor*. To do this we take into account that, for a deviatoric tensor, $\sigma_2 = -\sigma_1 - \sigma_3$, and obtain

$$\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & -\sigma_1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sigma_3 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}. \quad (18.36)$$

The two DCs have different orientations, the major DC with moment $M_0 = \sigma_1$ and the minor with $M_0 = \sigma_3$.

A more efficient separation is that proposed by Knopoff and Michael J. Randall (1970):

$$\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\sigma_1 - \sigma_3) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2}(\sigma_1 - \sigma_3) \end{pmatrix} + \begin{pmatrix} -\sigma_2/2 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & -\sigma_2/2 \end{pmatrix}. \quad (18.37)$$

As before, $\sigma_2 = -\sigma_1 - \sigma_3$. The first term is a DC source. The second is called a *compensated linear vector dipole* (CLVD). Its physical meaning is a sudden change in the shear modulus in a direction normal to the fault plane, without changes in volume. The source represented by a DC plus a CLVD corresponds to a shear fracture in which, during the rupture process, the shear modulus in the focal region changes suddenly. This separation represents the best solution that maximizes the DC part of the source.

In conclusion, a seismic point source of general type can be represented by the moment tensor. This source may involve changes in volume, shear fracture, and sudden changes in rigidity at the source, and thus can be separated in the form,

$$M = M^0 + M^{DC} + M^{CLVD}. \quad (18.38)$$

This partition separates shear fracture, considered the standard model for the source of earthquakes, from other effects that may also occur. The isotropic part is presupposed to be zero in many problems, but not in others such as volcanic earthquakes. A deviatoric source is formed by the DC plus CLVD sum. A deviation from a pure DC is sometimes represented by $\delta = |\sigma_3/\sigma_1|$, the ratio of the greatest and least eigenvalues. For a pure DC, $\delta = 1$, or more usually by

$$\varepsilon = \frac{\sigma_2}{\max(|\sigma_1|, |\sigma_3|)}, \quad (18.39)$$

and for a pure DC, $\varepsilon = 0$.

When the moment tensor is obtained from observations, the presence of non-DC components may be due to errors in observation or to propagation effects that have not been taken into account, rather than to the source itself. There is always a certain amount of ambiguity in distinguishing between effects that are due to the source and those due to propagation. Perfect separation of these two effects is not always possible.

18.5 Displacements due to a point source

According to (18.10), displacements due to a point source can be expressed by a time convolution of the moment tensor with derivatives of the Green function:

$$u_j(x_n, t) = \int_{-\infty}^{\infty} M_{kj}(\tau) G_{ikj}(t - \tau) d\tau. \quad (18.40)$$

Derivatives of the Green function for **P** and **S** waves in the far field for an infinite, homogeneous isotropic medium are given by (17.74) and (17.75). For **P** waves, time convolution according to (17.76) is given by

$$\int_{-\infty}^{\infty} M_{ij}(\tau) \delta\left(t - \frac{r}{\alpha} - \tau\right) d\tau = \dot{M}_{ij}\left(t - \frac{r}{\alpha}\right). \quad (18.41)$$

If we separate the modulus and time dependence from the orientation in the form $M_{ij}(t) = M_0(t)m_{ij}$, according to (17.77) and (17.78), we obtain the elastic displacements of **P** and **S** waves in the far field. In this expression we have assumed that all components of M_{ij} have the same time dependence (*synchronous moment tensor*) and, if it corresponds to a shear fracture, M_0 is the scalar seismic moment,

$$u_k^P = \frac{\dot{M}_0(t - r/\alpha)}{4\pi\rho\alpha^3 r} \gamma_i \gamma_j \gamma_k m_{ij}, \quad (18.42)$$

$$u_k^S = \frac{\dot{M}_0(t - r/\beta)}{4\pi\rho\beta^3 r} (\delta_{ik} - \gamma_i \gamma_k) \gamma_j m_{ij}. \quad (18.43)$$

Elastic displacements depend on the time derivative of the moment or the *moment rate*. Displacements for **P** waves and for **SV** and **SH** components of **S** waves are (calling the factors in (18.42) and (18.43) A and B),

$$u_P = A \gamma_i \gamma_j m_{ij}, \quad (18.44)$$

$$u_{SV} = -B SV_i \gamma_j m_{ij}, \quad (18.45)$$

$$u_{SH} = B SH_j \gamma_i m_{ij}, \quad (18.46)$$

where SV_i and SH_i are unit vectors in the **SV** and **SH** directions and we have taken into account that $SV_i \gamma_i = 0$ and $SH_i \gamma_i = 0$. Since γ_i is a unit vector in the ray's direction, **P** displacements are in the same direction and those of **S** are perpendicular (Section 5.7, Fig. 17.21). If the problem is referred to geographic axes (North, East, and nadir), for a homogeneous medium, γ_i , SV_i and SH_i can be given in terms of the azimuth ϕ' and take-off angle i of the ray:

$$\begin{aligned} \gamma_1 &= \sin i \cos \phi' \\ \gamma_2 &= \sin i \sin \phi' \\ \gamma_3 &= \cos i \\ SV_1 &= \cos i \cos \phi', \end{aligned} \quad (18.47)$$

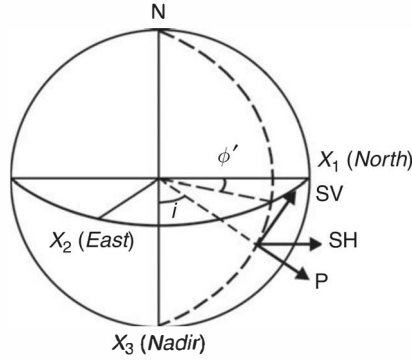


Fig. 18.6

Displacements of P, SH, and SV waves, and their relation to the angles i and ϕ' of the ray at the focus.

$$\begin{aligned} SV_2 &= \cos i \sin \phi' \\ SV_3 &= -\sin i, \end{aligned} \quad (18.48)$$

$$\begin{aligned} SH_1 &= -\sin \phi' \\ SH_2 &= \cos \phi' \\ SH_3 &= 0. \end{aligned} \quad (18.49)$$

The displacements u_P , u_{SV} , and u_{SH} can finally be given in terms of i and ϕ' (Fig. 18.6) and components of the moment tensor, by substituting equations (18.47)–(18.49) into equations (18.44)–(18.46):

$$\begin{aligned} u_P &= A \{ \sin^2 i [\cos^2 \phi' m_{11} + \sin^2 \phi' m_{22} + \sin(2\phi') m_{12}] \\ &\quad + \cos^2 i m_{33} + \sin(2i)(\cos \phi' m_{13} + \sin \phi' m_{23}) \}, \end{aligned} \quad (18.50)$$

$$\begin{aligned} u_{SV} &= B \left\{ \frac{1}{2} \sin(2i) [\cos^2 \phi' m_{11} + \sin^2 \phi' m_{22} - m_{33} + \sin(2\phi') m_{12}] \right. \\ &\quad \left. + \cos(2i)(\cos \phi' m_{13} + \sin \phi' m_{23}) \right\}, \end{aligned} \quad (18.51)$$

$$\begin{aligned} u_{SH} &= B \left\{ \sin i \left[\frac{1}{2} \sin(2\phi') m_{22} - \frac{1}{2} \sin(2\phi') m_{11} + \cos(2\phi') m_{12} \right] \right. \\ &\quad \left. + \cos i (\cos \phi' m_{23} - \sin \phi' m_{13}) \right\}. \end{aligned} \quad (18.52)$$

These three equations are for a general form of the moment tensor without assuming any particular condition.

18.6 The temporal dependence

For a point source moment tensor, if all of its components have the same time dependence (synchronous), a common assumption, the source time function, assuming a shear fracture,

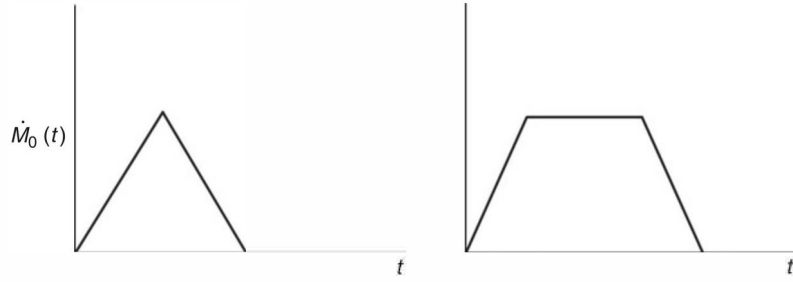


Fig. 18.7

The temporal dependence of $\dot{M}_0(t)$ for a simple source.

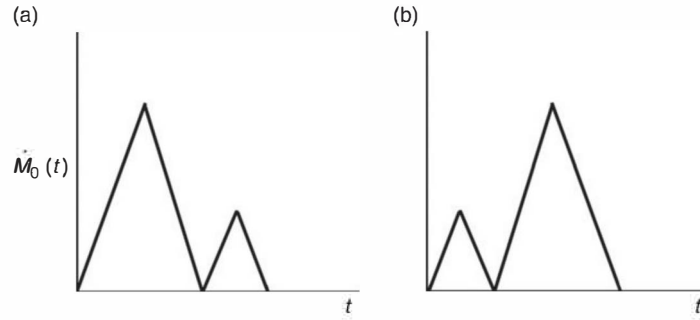


Fig. 18.8

The temporal dependence of $\dot{M}_0(t)$ for a complex source. (a) The main moment release occurs at the beginning. (b) The main moment release occurs at the end.

is given by $M_0(t)$. As we saw in (18.42) and (18.43), the displacements depend on the moment rate $\dot{M}(t)$, and its time dependence is also called the STF. This function represents the form according to which the moment rate changes with time, and its integral is the scalar seismic moment M_0 .

Just like for the time dependence of the slip rate, for a simple source a commonly used function for $\dot{M}_0(t)$ is a triangle or a trapezoid (Fig. 18.7). We have discussed the properties of this STF in Section 17.7 (Fig. 17.22). Since the size of an earthquake is given by M_0 , the same size will result in two different functions of the moment rate, one with a greater modulus and shorter duration and another with a smaller modulus and longer duration. We must remember that this is a point source representation, whereby the duration of the moment rate function can be related to an extended source with a specific duration of the fracturing process and thus to its dimensions. For a complex source, the moment rate may be represented by several triangles of different size (Fig. 18.8). The relative sizes of these sources show how the moment's radiation with time or *moment release* takes place. In some cases, the greatest part of the moment release occurs during the first part of the shock and a minor part follows later (Fig. 18.8a). Another possibility is that there is first a small moment release and

then, later, the main part (Fig. 18.8b). The small event may be considered as an aftershock in the first case and as a foreshock in the second (Section 21.2). Since these events are part of the fracturing process, the total duration of the source includes the complete moment release.

18.7 The centroid moment tensor

In the formulation of a point source in terms of the moment tensor we assume that this is acting at a point. This point is usually assumed to be the hypocenter. As defined in Section 16.1, the hypocenter location and origin time are obtained from the first arrivals of P and S waves of relatively short period, and thus correspond to the time and location of the beginning of rupture. Since rupture takes place on a finite size area the point of initiation of rupture may not be actually the best point source location and origin time to represent the source point. Dziewonski *et al.* (1981) introduced the source location and origin time as variables in their method of estimating the components of the moment tensor from observations of seismic waves. Thus they determine the point source location and origin time that do not correspond to the point and time of the initiation of rupture, but to a point that represents the *centroid* of the seismic moment tensor distribution in space and time. The difference of this point and that of the initiation of rupture can be significant except for very small earthquakes.

The position of the centroid can then be considered as a weighted mean position of the source or the geometrical center of gravity of the fault surface with respect to the moment-rate distribution. We saw (Section 16.1.1) that the macroseismic epicenter location may approximate to this point. The difference between the hypocenter and origin time and the centroid location and time is related to the dimensions and duration of the rupture process. The moment tensor located at the centroid is known as the *centroid moment tensor* (CMT). For the general case, a synchronous point source of this type is defined by twelve parameters (four for the location in space and time of the centroid (ϕ , λ , h , t), two for the scalar moment and rise time (M_0 , τ), and six for the normalized components of the moment tensor (m_{11} , m_{22} , m_{33} , m_{12} , m_{13} , m_{23})).

18.8 Inversion of the moment tensor

According to equations (18.8) and (18.9), for a point source (18.10), the elastic displacements are linear functions of the components of the moment tensor and derivatives of Green functions. In an explicit form, this is given in equations (18.50)–(18.52) for P and S waves in the far field of an infinite, homogeneous isotropic medium. This linear dependence makes it possible to determine the six components of the moment tensor from observations of the elastic displacements by linear inversion.

For a point source, the elastic displacements in the far field can be expressed in the time domain as a convolution (18.40):

$$u_k(t) = M_{ij}(t) * G_{kij}(t). \quad (18.53)$$

Taking the Fourier transform, in the frequency domain, the relation is a product of their transforms:

$$\bar{u}_k(\omega) = \bar{M}_{ij}(\omega) \bar{G}_{kij}(\omega). \quad (18.54)$$

In both cases, inversion consists of determination of the six components of M_{ij} from observed values of elastic waves u_j and components of appropriate functions for G_{kij} that are assumed to be known. Depending on the waves used (P, S, LR, or LQ) and the characteristics of the problem, the medium can be approximated by an infinite homogeneous medium, a half-space, or a layered medium using a flat or spherical Earth. Naturally, the Green functions become more complicated with increasing complexity of the Earth model. Derivatives of Green functions have in general 27 components, which for some particular cases can be reduced in number to eight.

If we impose no conditions, there are six components of the moment tensor to be determined. If the source is assumed to be purely deviatoric, $M_{11} + M_{22} + M_{33} = 0$, that is, $M_{33} = -M_{11} - M_{22}$, and the number of unknowns is five. This is a linear condition and the problem remains linear. If we assume the source to be a shear fracture or *DC*, the condition is that the determinant of M_{ij} is zero, and the problem ceases to be linear. For this reason this condition is not imposed for linear inversions.

For the linear problem, equations (18.53) and (18.54) may be written in matrix form as

$$\mathbf{u} = \mathbf{G}\mathbf{M}. \quad (18.55)$$

We need a number of observations larger than the number of unknowns (six or five), and solutions are obtained by a least squares procedure (see Section 16.1.3):

$$\mathbf{M} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{u}, \quad (18.56)$$

or, using the generalized inverse (16.7),

$$\mathbf{M} = \mathbf{U}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{u}, \quad (18.57)$$

where $\mathbf{U}$ and $\mathbf{V}$ are matrices formed by the eigenvectors of $\mathbf{G}^T \mathbf{G}$ and $\mathbf{G}\mathbf{G}^T$, respectively, and $\mathbf{A}$ is a diagonal matrix formed by the eigenvalues of $\mathbf{G}^T \mathbf{G}$. In practice, there are several methods for the solution of this problem (Chapter 20). Some methods use linear inversion whereas others use iterative procedures minimizing the differences between theoretical and observed displacements.

The observed data are the amplitudes of seismic waves recorded by seismographs. The problem can be solved in the time (18.53) or frequency (18.54) domain. In the frequency domain, the problem is linear either for the real or for the imaginary part of the spectrum, but not for the spectral amplitudes. In all cases, observations at different locations and at different times or frequencies are used so that the problem is well conditioned. Generally, we obtain the six components of the synchronous moment tensor, or five assuming that there is no change in volume. Since solutions do not necessarily correspond

to a DC source, the moment tensors obtained are separated into a DC and a non-DC part, as we saw in [Section 18.4](#). When the location and origin time are included in the inversion we obtain the centroid location and time and the CMT components. The main advantage of using the moment tensor formalism is that the source problem can be solved by linear inversion. Also we can represent the source in a general form and investigate whether there are volume changes and non-DC components. At present there are several agencies that routinely calculate the moment tensor for worldwide earthquakes above a certain magnitude ([Chapter 20](#)).

18.9 Summary

The seismic moment tensor is a very general and useful representation of the source mechanism of earthquakes. It represents the stresses which cause non-elastic displacements at the source. It is a symmetric tensor of which the elements of the diagonal are linear dipoles and the other couples with moments. Its eigenvalues are the principal stresses and its eigenvectors the axes of the principal stresses. It can be separated into an isotropic part which indicates the changes in volume and a deviatoric part with no volume changes. Because shear fracture is the most common type of source for earthquakes the moment tensor is separated into three parts: an isotropic part (M^0) with changes in volume, a part corresponding to a shear fracture or DC source (M^{DC}), and a non-DC deviatoric part, usually a CLVD (M^{CLVD}).

For a point source, the elastic field displacements are given by linear relations of the components of the moment tensor. For a synchronous moment tensor the STF is given by either $M_0(t)$ or $\dot{M}_0(t)$. A common STF used for $\dot{M}_0(t)$ is for a simple source, that of a triangular or trapezoidal shape, and for a complex one, a combination of triangles and trapezoids. The linear relation between the elastic displacements and the components of the moment tensor allows a linear inversion of the moment tensor from observations of seismic waves. If no conditions are imposed, the resulting moment tensor has parts corresponding to changes of volume and non-DC components. When the location and origin times are included in the inversion problem we obtain the centroid moment tensor.

18.10 Problems

- P18.1 The orientations of the T and P axes (Θ and Φ) are T (30° , 60°) and P (70° , 290°). Calculate the components of the moment tensor m_{ij} .
- P18.2 The components of the seismic moment tensor of an earthquake, with respect to the geographic coordinates (North, East, and down) are in units of 10^{18} Nm, $m_{11} = 2$, $m_{22} = 0$, $m_{33} = 2$, $m_{12} = -1$, $m_{13} = 1$, $m_{23} = 1$.
- (a) Find the eigenvalues and eigenvectors.

- (b) Separate it into isotropic and deviatoric parts.
 - (c) Separate the deviatoric part into major and minor DCs.
 - (d) Separate the deviatoric part into DC and CLVD parts. Find percentages of each.
 - (e) For the DC part of (d), find the scalar seismic moment M_0 and the orientations of the P and T axes (Φ, Θ). What type of fault is it?
- P18.3 An earthquake has a complex focal mechanism given by the sum of two DCs with orientation (ϕ, δ, λ) $(165^\circ, 89^\circ, 1^\circ)$ and $(122^\circ, 41^\circ, -130^\circ)$.
- (a) Find the components of the total seismic moment tensor.
 - (b) Separate it into isotropic and deviatoric parts.
 - (c) Separate the deviatoric part into DC and CLVD, estimating the percentage of each component.
 - (d) Estimate the orientation of fault plane (ϕ, δ, λ) corresponding to the DC component.
- P18.4 Calculate the components of the moment tensor corresponding to a shear fracture with $\phi = 45^\circ$, $\delta = 60^\circ$, $\lambda = -90^\circ$, and $M_0 = 10^{15}$ Nm. From these values calculate the displacements of P waves at a point at a distance $r = 500$ km, with $i_h = 15^\circ$ and $\alpha_z = 15^\circ$. The medium has $\alpha = 8$ km s⁻¹ and $\rho = 3$ g cm⁻³.

In [Chapters 17](#) and [18](#) we have considered in some detail the characteristics and elastic displacement field corresponding to point sources. A more complete representation of the seismic source must include its dimensions and fracture propagation, and consider their effects on wave radiation. First considerations of the dimensions of the seismic focus proposed models consisting of spherical cavities of finite radii with uniform distribution of stresses on their surface ([Jeffreys, 1931](#); [Nishimura, 1937](#); [Scholte, 1962](#)).

The first models for extended sources of shear fracture were *kinematic models*, consisting of a slip that propagates with a constant velocity over a surface of finite area. Ari Ben Menahem ([1961, 1962](#)) described extended sources in terms of distributions of single and double couples propagating with a certain velocity over a rectangular surface, and determined the corresponding displacements of body and surface waves. Hans Berckhemer ([1962](#)) studied the effect on the width of temporal pulses of a circular fracture of finite radius that propagates from its center. Robert Burridge and Knopoff ([1964](#)) treated shear dislocations that propagate over a certain area and showed their equivalence to propagating double couples. Haskell ([1964, 1966](#)) proposed a rectangular model of fracture, and Brian Savage ([1966](#)) proposed an elliptical fault and studied the effects on the spectra of body and surface waves. James Brune ([1970](#)) presented a model with shear stresses suddenly applied to a circular fault, and studied elastic displacements in near and far fields. More recent kinematic models include propagating shear fractures on finite faults with variable slip, rupture velocity, and rise time ([Hartzell, 1989](#)).

The physical problem of fracture is a dynamic problem in which the slip has to be considered a consequence of stress conditions and the strength of material in the focal region. *Dynamic models* of the seismic source take these conditions into consideration and are based on the theory of the generation and propagation of fractures in stressed media. A complete dynamic model must, then, include the whole of the fracturing process, its initiation or nucleation, propagation and arrest derived from stress conditions, and properties of the material in the focal region. Two determinant factors are tectonic stresses which are a consequence of lithospheric plate motion and mechanical properties of rocks in the fault region. Among the first studies of fracture dynamics applied to earthquakes were those of Keylis-Borok ([1959](#)), Boris V. Kostrov ([1964, 1966](#)), Burridge ([1969](#)), Lambert B. Freund ([1972, 1979](#)), and Raúl Madariaga ([1976](#)). These studies were based on the work on fractures of crystals and metals published between 1920 and 1950 by Alan A. Griffith, A. T. Starr, and George R. Irwin. The dynamic problem is more complicated than the kinematic problem; therefore, we will present only the more basic principles in a simplified form.

19.1 Source dimensions. Kinematic models

Let us consider first some general characteristics of kinematic models of extended sources represented by a surface Σ over which a shear dislocation $\Delta \mathbf{u}(\xi_i)$ propagates with a constant velocity v in one direction, from the origin ($\xi_i = 0$) to a final point over a distance L (Fig. 19.1) (Aki and Richards, 1980). The total time of rupture is given by $t_r = L/v$. The velocity of the fracture's propagation is assumed to be constant and less than the velocity of wave propagation ($v < \beta < \alpha$); that is, we treat subsonic fractures (a common value is $v = 0.7\beta$). From equations (17.71) and (17.77), the displacements of P waves in the far field for an infinite, homogeneous, isotropic medium can be written as

$$u_i^p(x_j, t) = \frac{\mu}{4\pi\alpha^3\rho} \int_{\Sigma} \frac{R(n_k, l_k, \gamma_k)}{r} \Delta \dot{u} \left(\xi_i, t - \frac{r}{\alpha} \right) dS, \quad (19.1)$$

where $r = |x_i - \xi_i|$ is the distance from the point of observation x_i to a point of the source ξ_i where the slip $\Delta \mathbf{u}$ is located at each moment, and $R(n_k, l_k, \gamma_k)$ is the radiation pattern that depends on the orientation of the source ($\mathbf{l}$, $\mathbf{n}$) and the position of the observation point (γ_i). If we are interested only in the wave form as a function of time at a certain distance r_0 from the origin of ξ_i (Fig. 19.1), we need only consider the integral (Section 17.3),

$$\mathbf{u}(t) = \int_{\Sigma} \Delta \dot{u} \left(\xi_i, t - \frac{r}{\alpha} \right) dS. \quad (19.2)$$

Expressing r in terms of a Taylor expansion about r_0 , we obtain

$$r = r_0 + \xi_i \frac{\partial r}{\partial \xi_i} + \frac{1}{2} \xi_i^2 \frac{\partial^2 r}{\partial \xi_i^2} + K, \quad (19.3)$$

and, since (17.53) applies,

$$\frac{\partial r}{\partial \xi_i} = -\gamma_i. \quad (19.4)$$

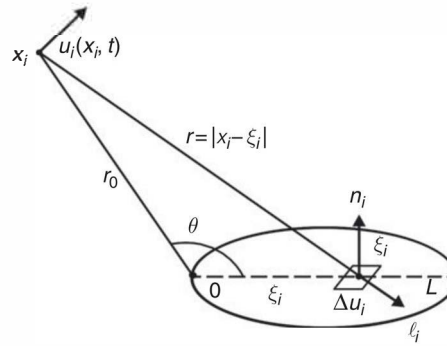


Fig. 19.1

An extended source of dimension L , with slip $\Delta \mathbf{u}$ at ξ_i and an elastic displacement $\mathbf{u}$ at x_i .

If we keep only the first-order term of ξ_i and neglect those involving higher powers, we obtain

$$r = r_0 - \xi_i \gamma_i. \quad (19.5)$$

Since the first neglected term is ξ_i^2/r_0 , and the maximum value of $|\xi_i|$ is L , the approximation given by (19.5) is good for displacements of wave length λ greater than the neglected term, that is, for $\lambda r_0 \gg L^2$. For these wave lengths, elastic displacements of P waves are given by (19.2):

$$u(t) = \int_{\Sigma} \Delta \dot{u} \left(\xi_i, t - \frac{r_0 - \xi_i \gamma_i}{\alpha} \right) dS. \quad (19.6)$$

Its Fourier transform (Appendix 4) is

$$U(\omega) = \int_{\Sigma} \Delta \dot{U}(\omega) e^{-i\omega(r_0 - \xi_i \gamma_i)/\alpha} dS, \quad (19.7)$$

where $\Delta \dot{U}(\omega)$ is the FT of $\Delta \dot{u}(t)$ and we have used that the FT of $f(t-a) = F(\omega) \exp(-i\omega a)$. Since $\Delta \dot{U}(\omega) = i\omega \Delta U(\omega)$, we finally obtain for $U(\omega)$,

$$U(\omega) = e^{-i\omega r_0/\alpha} \int_{\Sigma} i\omega \Delta U(\omega) e^{i\omega \gamma_i \xi_i/\alpha} dS. \quad (19.8)$$

Thus, the transform of elastic displacements $U(\omega)$ has the form of a spatial transform over the fault plane of the transform of the slip $\Delta U(\omega)$. In the exponential, the factor $\omega \gamma_i/\alpha = k \gamma_i$ is the projection of the wave number k onto the fault plane Σ . If the slip has a step function time dependence $\Delta u(t, \xi_i) = \Delta u(\xi_i) H(t)$, its transform is $\Delta u(\xi_i)/(i\omega)$. By substituting into (19.8), we obtain

$$U(\omega) = e^{-i\omega r_0/\alpha} \int_{\Sigma} \Delta u(\xi_i) e^{i\omega \gamma_i \xi_i/\alpha} dS. \quad (19.9)$$

If we take the limit for low frequencies, then, when ω tends to zero, introducing the factor μ from (19.1), we obtain

$$U(0) \simeq \mu \int_{\Sigma} \Delta u(\xi_i) dS \simeq \mu \Delta \bar{u} S \simeq M_0. \quad (19.10)$$

For low frequencies, the spectral amplitudes are proportional to the seismic moment. The proportionality depends on the factors present in (19.1). For high frequencies, as they tend to infinity, if the slip does not change sign, the spectral amplitudes tend to zero. The form in which $U(\omega)$ tends to zero from the limit of $U(0)$ depends on the form of the slip function.

In general, the form of amplitude spectra corresponding to a source with finite dimensions is the following: $U(\omega)$ is constant for a range of low frequencies and starts to decrease from a certain frequency that, as will be shown later, is proportional to the inverse of the source dimensions. The envelope of the spectrum in the high frequencies has a frequency dependence of $\omega^{-\varepsilon}$, where ε has values between zero and three, and generally equals two (Aki, 1967). Thus, high-frequency spectra are limited to a certain range of frequencies. For point

sources and a step source time function, we saw that the far-field elastic displacements of P and S waves are impulses (delta functions) ((17.79) and (17.80)) and in consequence their spectra are constant for all frequencies (the FT of the delta function is a constant). The source dimensions limit the high-frequency spectra to a certain maximum value. The larger the source dimension the lower the frequencies present in the spectrum. These results have been obtained without specifying the form of the source and are common characteristics for all sources of finite dimensions such that the slip at each point starts at zero and ends with a constant value.

19.2 Rectangular faults. Haskell's model

A simple kinematic model of finite dimensions, known as *Haskell's model*, is a rectangular fault of length L and width W , such that the slip Δu propagates only along the L direction with a constant velocity v (the slip moves instantaneously along W) (Haskell, 1964). The coordinate along L is ξ , with its origin at one end of the fault and Δu has only one component (Fig. 19.2). Fractures that propagate only in one sense (from 0 to L) are called *unilateral fractures* and those that propagate in both senses (from 0 to $L/2$ and from 0 to $-L/2$) are called *bilateral fractures*. For unilateral fractures, according to (19.6), not considering the radiation pattern, the dependence on distance, and the other factors of (19.1), the form of P waves in the far field is given by

$$u(x_i, t) = W \int_0^L \Delta u \left(\xi, t - \frac{r_0 - \xi \cos \theta}{\alpha} \right) d\xi. \quad (19.11)$$

If the slip moves in the positive direction of ξ with a constant fracture velocity v , then $\Delta u(\xi, t) = \Delta u(t - \xi/v)$ and we obtain

$$u(x_i, t) = W \int_0^L \Delta u \left[t - \frac{r_0}{\alpha} - \frac{\xi}{\alpha} \left(\frac{\alpha}{v} - \cos \theta \right) \right] d\xi. \quad (19.12)$$

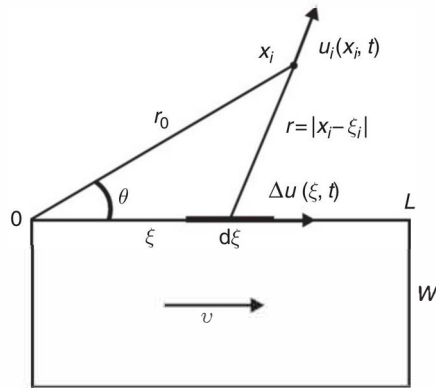


Fig. 19.2

The rectangular fault of Haskell's model.

If we make the substitution,

$$d = \frac{r_0}{\alpha} + \frac{\xi}{\alpha} \left(\frac{\alpha}{v} - \cos \theta \right), \quad (19.13)$$

then the Fourier transform of $u(x_i, t)$ is

$$U(x_i, \omega) = W \int_0^L d\xi \int_{-\infty}^{\infty} \Delta \dot{u}(t-d) e^{-i\omega(t-d)} dt. \quad (19.14)$$

However, we have that

$$\int_{-\infty}^{\infty} \Delta \dot{u}(t-d) e^{-i\omega(t-d)} dt = i\omega \Delta U(\omega) e^{-i\omega d}, \quad (19.15)$$

where, as before ((19.7) and 19.8)) $\Delta U(\omega)$ is the transform of $\Delta u(t)$, the transform of $\Delta \dot{u}(t)$ is $i\omega \Delta U(\omega)$, and that of $\Delta u(t-d) = \Delta U(\omega) \exp(-i\omega d)$. Therefore, equation (19.14) becomes

$$U(x_i, \omega) = W i\omega \Delta U(\omega) e^{-i\omega r_0/\alpha} \int_0^L \exp \left[-i \frac{\xi\omega}{\alpha} \left(\frac{\alpha}{v} - \cos \theta \right) \right] d\xi. \quad (19.16)$$

To evaluate the integral in (19.16), we make the substitution $b = -(\omega/\alpha)(\alpha/v - \cos \theta)$ and obtain

$$\int_0^L e^{ib\xi} d\xi = \frac{2}{b} \sin \left(\frac{bL}{2} \right) \exp \left(i \frac{bL}{2} \right) = L \frac{\sin X}{X} e^{iX}, \quad (19.17)$$

where

$$X = \frac{bL}{2} = -\frac{\omega L}{2\alpha} \left(\frac{\alpha}{v} - \cos \theta \right). \quad (19.18)$$

The final form for the transform of elastic displacements of P waves $U(x_i, \omega)$, according to (19.16), is

$$U(x_i, \omega) = WL\omega \Delta U(\omega) \frac{\sin X}{X} \exp \left[-i \left(\frac{\omega r_0}{\alpha} - X - \frac{\pi}{2} \right) \right], \quad (19.19)$$

where we have replaced $i = e^{i\pi/2}$. The form of the amplitude spectrum depends on the factor $\sin X/X$. We have discussed the form of this function in Section 13.2. It has the value unity for $X = 0$ and roots for X equal to integer multiples of π , and its envelope decreases as $1/X$ (Fig. 19.3). Since, for fixed values of θ and L , X depends on ω , in the limit when ω tends to zero (low frequencies), the factor equals unity and for high frequencies its envelope decreases with $1/\omega$.

The form of the amplitude spectrum depends also on the form of $\Delta U(\omega)$, the transform of the STF (Section 17.7) (19.19). If $\Delta u(t) = \Delta u H(t)$ (instantaneous slip), its transform is $\Delta U(\omega) = \Delta u/(i\omega)$. From (19.19) we obtain that $U(\omega)$ is proportional to the seismic moment ($M_0 = \mu L W \Delta u$) for the limit of low frequencies (as we saw in (19.10) and decreases as $1/\omega$ for high frequencies. If the STF has a rise time τ that Δu takes to attain its maximum value at each point of the fault plane (Section 17.7), the spectrum depends on

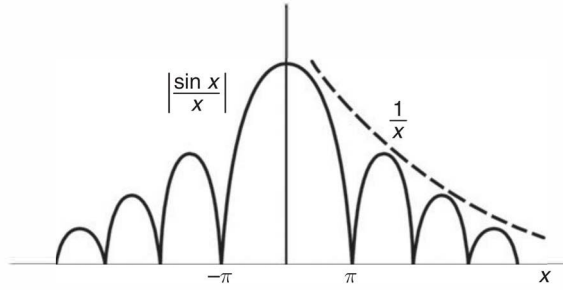


Fig. 19.3 The function $\sin X/X$.

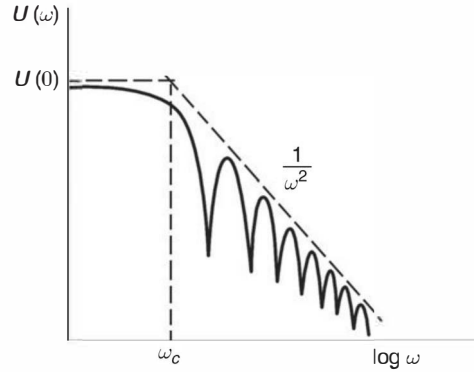


Fig. 19.4 The form of the amplitude spectrum of seismic waves for an extended fault with finite dimensions and rise time.

the transform of the STF. For example, the transforms of the STF given by (17.109) and (17.110) are

$$\Delta u(t) = \begin{cases} \Delta u t / \tau, & 0 < t < \tau \\ \Delta u, & t \geq \tau \end{cases}; \quad \Delta U(\omega) = \frac{\Delta u(1 - e^{-i\omega\tau})}{\omega^2 \tau}, \quad (19.20)$$

$$\Delta u(t) = \Delta u H(t)(1 - e^{-t/\tau}); \quad \Delta U(\omega) = \frac{\Delta u}{(1 + i\omega\tau)i\omega}. \quad (19.21)$$

In both cases, the transforms depend on $1/\omega^2$. If we substitute these values of $\Delta U(\omega)$ into (19.19), the envelope of $U(\omega)$ decreases with the frequency as $1/\omega^2$. If we represent the spectrum with respect to the logarithm of the frequency, its form is a flat part for low frequencies and, from a certain frequency ω_c , called the *corner frequency*, its envelope is a straight line of slope -2 (Fig. 19.4). This form of the spectrum is due to the combined effect of the source dimensions and the rise time. If we consider the particular case in which $\theta = \pi/2$ and ω_c corresponds to $X = \pi/2$, we obtain $\omega_c = 2v/L$; that is, the corner frequency is proportional to the inverse of the source length. Observed spectra of seismic waves exhibit these characteristics, indicating the combined effect of the finite dimensions of the source and the existence of the rise time (Aki, 1967).

The influence of the source dimensions can be isolated by means of the *directivity function* $D(\omega)$, defined by Ben Menahem (1961) as the quotient of spectral amplitudes of waves that leave the source in opposite directions, that is, with angles θ and $\theta + \pi$. According to (19.18) and (19.19), this quotient for a vertical fault is

$$D(\omega) = \frac{|\sin\{[\omega L/(2c)](c/v - \cos\theta)\}|(c/v + \cos\theta)|}{|\sin\{[\omega L/(2c)](c/v + \cos\theta)\}|(c/v - \cos\theta)|}, \quad (19.22)$$

where c is the wave velocity. This function has a series of maxima and minima for frequencies that depend on L and v , and can be used to determine the source dimensions, velocity, and sense of fracture propagation. This is easier for surface waves, since θ represents the azimuth at the focus with respect to the trace of the fault.

Another effect of equation (19.19) is on the form of the radiation pattern (Section 17.6.1 and Fig. 17.18). If the wave length is much larger than the source dimensions, $\lambda \gg L$, X tends to zero and $\sin X/X$ is unity for all values of θ . Amplitudes are not affected and the radiation pattern corresponds to that of a point source. If the wave length is of the same order as the dimensions ($\lambda \approx L$), amplitudes are affected by the factor $\sin X/X$ which depends on θ , and the radiation pattern is modified. According to (19.18), this factor is maximum for $\theta = 0$ and minimum for $\theta = \pi$; that is, amplitudes are larger in the same direction as that of fracture propagation ($\theta = 0$) and smaller in the opposite direction ($\theta = \pi$) (Fig. 19.5). This effect is called the *focusing of energy* in the direction of fracture propagation and is a phenomenon that occurs for all wave sources which propagate along a finite distance.

The kinematic model of a rectangular unilateral fracture with a constant rupture velocity has shown us the effects of the source dimensions on the radiated displacement field. Amplitude spectra of displacements have constant values proportional to the seismic moment at low frequencies and these values decrease with frequency for high frequencies, starting from the corner frequency. If the STF includes a rise time, this decrease corresponds to $1/\omega^2$. The radiation pattern is also affected by dimensions, with more energy radiated in the azimuth corresponding to the direction of propagation of the rupture.

Haskell's model with bilateral fracture, a rupture velocity of $v = 0.9\beta$, and a STF given by (19.21) has two corner frequencies ω_1 and ω_2 instead of one (Savage, 1972). For

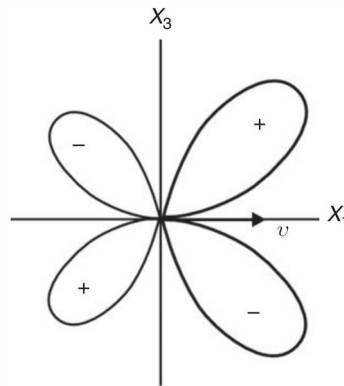


Fig. 19.5

The effect of fracture propagation on the radiation pattern of P waves.

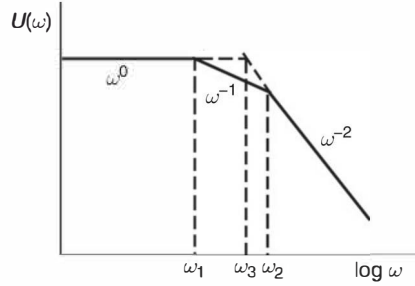


Fig. 19.6 The amplitude spectrum of seismic waves according to Savage's model.

frequencies between zero and ω_1 , the spectrum is flat; between ω_1 and ω_2 it decreases as ω^{-1} ; and for frequencies higher than ω_2 , it decreases as ω^{-2} (Fig. 19.6). A third corner frequency ω_3 is defined by the intersection of the flat part and the decay as ω^{-2} . For P and S waves ω_1 , ω_2 , and ω_3 are given by

$$\begin{aligned}
 \text{P : } \omega_1 &= \frac{1.2\alpha}{L} & \text{S : } \omega_1 &= \frac{3.6\beta}{L} \\
 \omega_2 &= \frac{2.4\alpha}{W} & \omega_2 &= \frac{4.1\beta}{W} \\
 \omega_3^2 &= \frac{2.9\alpha^2}{LW} & \omega_3^2 &= \frac{14.8\beta^2}{LW}
 \end{aligned} \tag{19.23}$$

The corner frequencies of P waves are always lower than those of S waves. Usually, observed corner frequencies ω_c correspond to ω_3 , and from this value we can obtain the source dimensions:

$$(LW)^{1/2} = \frac{1.7\alpha}{\omega_c^P} = \frac{3.8\beta}{\omega_c^S}. \tag{19.24}$$

The difference between ω_1 and ω_2 depends on the relation between L and W . If $W \ll L$, that is, the fault is long and narrow, then the difference is large, whereas, if $L \approx W$, the three frequencies practically coincide.

19.3 Circular faults. Brune's model

Another fundamental model of an extended seismic source is that of a circular fault, known as *Brune's model* (Brune, 1970). This model consists of a circular fault plane with finite radius on which a shear stress pulse is applied instantaneously (Fig. 19.7). Since this model specifies the stress on the fault, this is not exactly a kinematic model. Because the stress pulse is applied instantaneously on the whole fault area, there is no fracture propagation. The shear pulse generates a shear wave that propagates perpendicularly to the fault plane. Adapting Brune's notation to that we have used, we call $\Delta\sigma$ Brune's *effective shear stress*

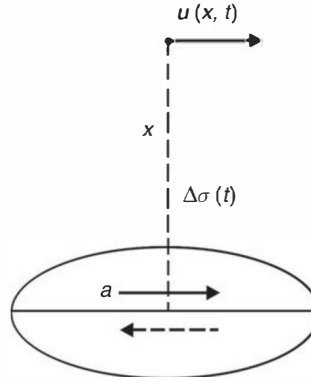


Fig. 19.7 Brune's model of a circular fault.

and Δu the displacement on the fault plane (that is, for $x = 0$, where x is the distance normal to the fault plane). The stress pulse has a time dependence given by a step function and, for a distance x , is

$$\Delta\sigma(x, t) = \Delta\sigma H\left(t - \frac{x}{\beta}\right). \quad (19.25)$$

The shear displacement Δu , for $x = 0$, is obtained by integration of (19.25), since, in this case, $\sigma = \mu \partial u / \partial x$:

$$\Delta u(t) = H(t) \frac{\Delta\sigma}{\mu} \beta t. \quad (19.26)$$

Its Fourier transform is

$$\Delta U(\omega) = -\frac{\Delta\sigma \beta}{\mu \omega^2}. \quad (19.27)$$

The effective stress is the difference between the tectonic stress σ_0 acting on the fault plane before faulting and the friction stress after it σ_f ($\Delta\sigma = \sigma_0 - \sigma_f = \varepsilon\sigma_0$). This coincides with the drop in stress defined in (16.34). For a total drop in stress ($\Delta\sigma = \sigma_0$ and $\varepsilon = 1$) the displacement of S waves in the far field at distance r , not including the radiation pattern and the dependence on distance, is

$$u(t) = \frac{\Delta\sigma\beta}{\mu} \left(t - \frac{r}{\beta}\right) \exp\left[-b\left(t - \frac{r}{\beta}\right)\right], \quad (19.28)$$

where

$$b = \frac{2.33\beta}{a}. \quad (19.29)$$

Its spectrum is

$$U(\omega) = \frac{\Delta\sigma\beta}{\mu} \frac{1}{\omega^2 + b^2}, \quad (19.30)$$

where a is the radius of the fault. The spectrum (19.30) has a flat part at low frequencies as they tend to zero and decreases as ω^{-2} for high frequencies, starting at the corner frequency $\omega_c = b$. If the stress drop is not total, then, for small ε ($\varepsilon \approx 0.01$), the spectrum decreases as ω^{-1} . The fault radius can be deduced according to (19.29) from the corner frequency ω_c of S waves:

$$a = 2.33\beta/\omega_c. \quad (19.31)$$

Brune's model is commonly used to obtain fault dimensions from spectra of S waves for earthquakes of small to moderate size ($M < 6$), for which the circular fault is a good approximation. We have mentioned (Section 17.1) that earthquakes take place in the brittle part of the crust or the seismogenic layer. For dimensions less than 20 km ($M < 6$), fault planes are contained inside the seismogenic layer. Fractures start at a point and grow unhindered in all directions with near circular form ($L \approx W$) and can be approximated by Brune's model. Larger earthquakes have larger dimensions, so, since their width is limited to about 20 km, their length must be larger than their width ($L > W$). In this case, Haskell's rectangular model is a better approximation.

19.4 Nucleation, propagation, and arrest of a rupture

Haskell's model does not include the effect either of the beginning or *nucleation* of a rupture or of its cessation or *arrest*. The first kinematic model that included both effects was proposed by Savage (1966). The Savage model consists of an elliptical fault in which slip begins at one of the foci and stops when it reaches the border of the ellipse. The model can be simplified for a circular fault of radius a , for which the slip Δu (which is constant for all points) begins at the center, propagates radially with a constant rupture velocity v and circular rupture fronts, and stops at the circular border. We use polar coordinates (ρ, ϕ) on the fault plane, ρ , with its origin at the center of the fault (not to be confused with the density) and ϕ measured from x_1 . Then, the time dependence of the slip is a step function and the slip is a function only of ρ given in the form,

$$\Delta u(\rho, t) = \Delta u H\left(t - \frac{\rho}{v}\right) [1 - H(\rho - a)]. \quad (19.32)$$

At the center of the fault ($\rho = 0$), the slip at $t = 0$ passes instantaneously from zero to Δu . For a value at a distance ρ ($0 < \rho < a$) from the center, the slip is 0 until $t = \rho/v$, when it becomes Δu . For $\rho = a$, the rupture stops and, for $\rho \geq a$, the slip is zero for all values of t (Fig. 19.8).

For a point of observation on the x_3 axis, that is, over the center of the fault at a distance r_0 (Fig. 19.9a), the form of P waves, according to (19.2), is given by

$$u(r_0, t) = \int_0^{2\pi} \int_0^a \Delta \dot{u}\left(\rho, t - \frac{r}{\alpha}\right) \rho \, d\rho d\phi. \quad (19.33)$$

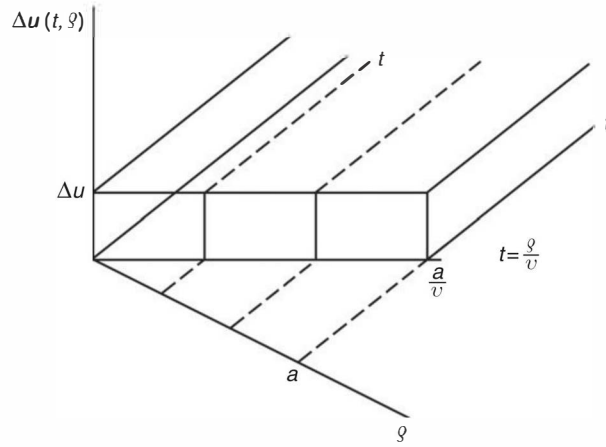


Fig. 19.8

The slip as a function of time and distance inside the fault for the Savage model of initiation and cessation of rupturing.

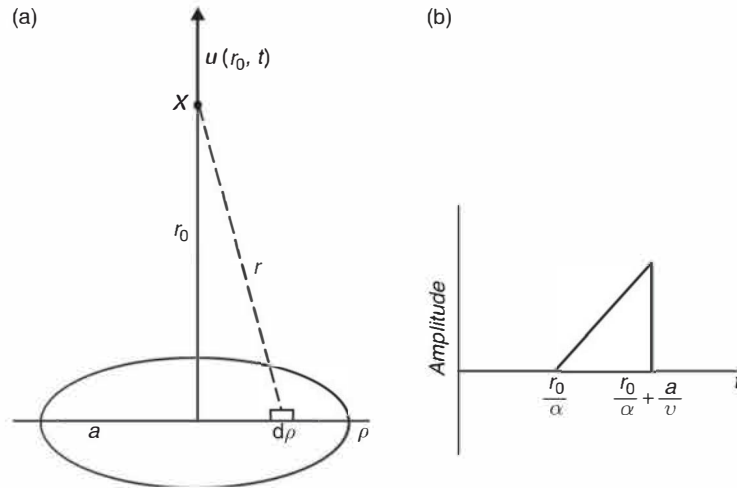


Fig. 19.9

Effects of initiation and cessation of a circular fault for points on a perpendicular axis at its center. (a) The fault and observation point. (b) The elastic displacement of P waves at a distance r_0 along the perpendicular axis.

Since rupture propagates in the ρ direction with a velocity v , as in (19.12), the slip can be written as $\Delta u(t - r/\alpha - \rho/v)$. Substituting into (19.33) and with the approximation that, for $r_0 \gg a$, $r = r_0$, after integration over ϕ , we have

$$u(r_0, t) = 2\pi \int_0^a \Delta \dot{u} \left(t - \frac{r_0}{\alpha} - \frac{\rho}{v} \right) \rho \, d\rho. \quad (19.34)$$

Taking the time derivative in (19.32) and substituting it into (19.34), the displacement is given by

$$u(r_0, t) = 2\pi \Delta u H\left(t - \frac{r_0}{\alpha}\right) \int_0^a \delta\left(t - \frac{\rho}{v} - \frac{r_0}{\alpha}\right) [1 - H(\rho - a)] \rho \, d\rho. \quad (19.35)$$

To evaluate this integral we use the relation

$$\int_{-\infty}^{\infty} f(x) \delta(ax - b) \, dx = \frac{1}{a} f\left(\frac{b}{a}\right). \quad (19.36)$$

The integral of (19.35) can be extended to the interval $(-\infty, \infty)$ since it is zero outside the interval $(0, a)$, and thus, by applying (19.36), we obtain

$$u(r_0, t) = 2\pi \Delta u v^2 H\left(t - \frac{r_0}{\alpha}\right) \left(t - \frac{r_0}{\alpha}\right) \left\{1 - H\left[v\left(t - \frac{r_0}{\alpha}\right) - a\right]\right\}. \quad (19.37)$$

According to this expression, for

$$t \leq \frac{r_0}{\alpha} \quad \text{and} \quad t \geq \frac{r_0}{\alpha} + \frac{a}{v}, \quad u(r_0, t) = 0, \quad (19.38)$$

whereas for

$$\frac{r_0}{\alpha} < t < \frac{r_0}{\alpha} + \frac{a}{v}, \quad u(r_0, t) = 2\pi \Delta u v^2 \left(t - \frac{r_0}{\alpha}\right). \quad (19.39)$$

Displacement starts at $t = r_0/\alpha$ and increases linearly with time until $t = a/v + r_0/\alpha$, when it drops to zero (Fig. 19.9b). According to the approximation used ($r = r_\bullet$), the time of inception of the discontinuity corresponds to the arrival of the signal from the cessation of the fracture at the border ($\rho = a$), which is called the *stopping phase*. The displacement drops discontinuously to zero and the velocity and acceleration become infinite.

In the Savage model the slip passes instantaneously from zero to its maximum value at each point of the fault as the rupture propagates from the center outward. The slip velocity is a pulse that propagates in the same way until it reaches the border of the fault at a time $t_r = a/v$. Since elastic displacements depend on the slip velocity, other models specify this value directly (as was done for the STF in Section 17.7) (Molnar *et al.*, 1973). For a circular fault the slip velocity can be expressed as

$$\Delta \dot{u}(\rho, t) = \Delta V_M \left[H\left(t - \frac{\rho}{v}\right) - H\left(t + \frac{\rho}{v} - \frac{a}{v}\right) \right] H(a - \rho). \quad (19.40)$$

In this model, the slip velocity takes a constant maximum value $\Delta \dot{u} = \Delta V_M$ at each point of the fault as the rupture front arrives at $t = \rho/v$, which remains constant as long as $t < (a - \rho)/v$ and ceases for $t \geq (a - \rho)/v$. This means that the slip velocity persists at each point of the fault until fracture stops at the border $\rho = a$ at a time t_r (Fig. 19.10). However, it is not physically possible to stop motion in the fault in this way, since this implies that information on the cessation of the fault at the border *instantaneously* (with infinite velocity) reaches all points of the fault. This can be solved by introducing a finite velocity, for example, the velocity of P waves, to bring the information on the cessation of rupture at the border to all points of the interior of the fault. This can be done by replacing the second

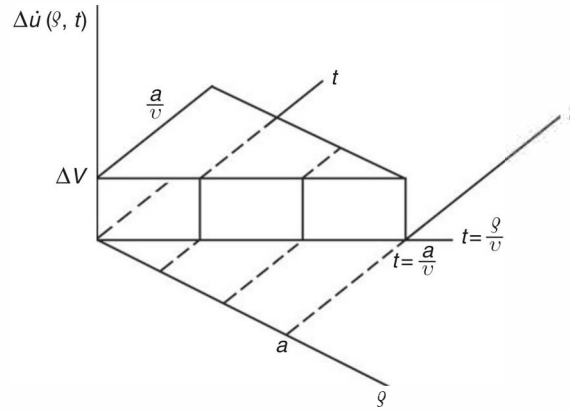


Fig. 19.10

The source model for the slip velocity in a fault with initiation and cessation.

term inside the square brackets of (19.40) by $H[t - a/v - (a - \rho)/a]$. Now the slip velocity becomes zero at each point of the fault as a P wave arrives from the fault border, once the motion has stopped at the border. This wave is called the *healing front*, since it heals the fault by stopping its motion. In kinematic models the healing of the rupture inside the fault must be introduced in some way. Models represented by (19.32) and (19.40) have the slip and slip velocity time dependence of a step function. Just like for point sources, we can also introduce a rise time into these models (Section 17.7). As the rupture front reaches each point of the interior of the fault, the slip velocity starts to increase from zero to a maximum value during a time τ (the rise time). Finally, the slip velocity $\Delta\dot{u}(t)$ is brought down to zero at each point of the fault as a healing front arrives from the border of the fault where it has stopped.

Another way to stop motion at each point of the fault is by introducing that at each point, slip velocity begins when $t = \rho/v$ reaches a maximum and decreases to zero after a certain rise time τ (Section 17.7). In this case there is no need for a healing front, since motion ceases at each point of the fault after a fixed time τ . As in most kinematic models the rise time τ and the total duration of fracture t_r are independent; we must consider the relation between them. If both are equal ($\tau \approx t_r$) motion at each point ceases as the same time as the total time of fracture. Most commonly τ is assumed to be much smaller than t_r and then $\Delta\dot{u}(t)$ propagates over the fault plane as a pulse, called a *Heaton pulse* after Thomas H. Heaton, healing itself after time τ at each point (Heaton, 1990). More realistic kinematic models can be established with various shapes, generally rectangular, in which the slip and slip velocities decrease gradually with rupture near to the border of the fault, and the fracture velocity, maximum slip, slip velocity, and rise time vary along the fault plane (Archuleta, 1984; Mendoza and Hartzell, 1989). In general, kinematic models of faulting can be made to correspond quite realistically with conditions on a fault, but they are not completely exempt from a certain arbitrariness. Some conditions must be imposed on the faulting process *a priori*, such as the velocity of rupture propagation, stopping at the border, and the healing process.

19.5 Dynamic models of fracture

The kinematic models we have considered up to this point are, naturally, simplifications of real fractures and include certain arbitrary factors in the definition of the slip and conditions at the fault border. The dynamic model relates the slip on the fault with the drop in stress that produces it. On the fault, the fracture front propagates with a certain velocity and as it advances, material becomes fractured and stress drops to a residual value. From this point of view, the mechanism of an earthquake is represented by a shear fracture produced by the drop in stress in the focal region. Fracture initiates at a point of the fault when the stress acting on the fault plane exceeds a critical value, propagates with a certain velocity, and finally stops when conditions impede its further propagation. Let us consider the energy produced by the fracture process given by (16.36). From the dynamic point of view, a fracture is produced by a drop in stress $\Delta\sigma = \sigma_0 - \sigma_f$, where σ_0 is the shear component of the sum of the tectonic stresses acting before faulting and σ_f is the friction between the two sides of the fault. By substitution into (16.36), assuming that the average stress $\bar{\sigma} = \Delta\sigma/2$, we obtain

$$E = \frac{1}{2} \Delta\sigma \Delta u S + \sigma_f \Delta u S, \quad (19.41)$$

where the first term represents the seismic energy due to the slip on the fault and the second represents the residual energy lost by friction. From the dynamic point of view, the stress drop is the parameter that determines the fracture and the slip is its consequence.

19.5.1 The static problem

Let us start with the static problem of a fracture free from stress after faulting. For a circular fault of radius a , the shear stress before faulting is σ_0 and, after faulting, the stress inside the fault using polar coordinates ($\rho < a$) is null. Then, $\Delta\sigma_s = \sigma_0$ is the static drop in stress. In the absence of body forces ($\mathbf{F} = 0$), for the static case, the equation of motion (4.65) is $\tau_{ij,j} = 0$. For shear slip on the fault $u(\rho)$, in a homogeneous isotropic elastic medium, the equation becomes

$$\mu \nabla^2 u = 0. \quad (19.42)$$

The conditions of displacement and stress inside and outside the fault are

$$\begin{array}{l|l} \rho = 0, & u(0) = \Delta u \\ \rho \geq a, & u(\rho) = 0 \\ \rho < a, & \sigma(\rho) = 0 \\ \rho = \infty, & \sigma(\rho) = \sigma_0 \end{array}$$

Under these conditions, the solution for the displacement inside the fault is

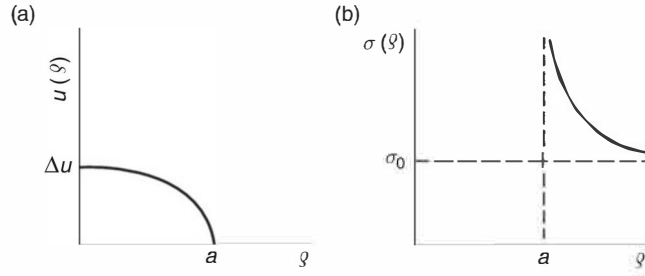


Fig. 19.11 The static model of fracture with a total drop in stress: (a) slip and (b) stress.

$$u(\rho) = \frac{\sigma_0}{\mu} (a^2 - \rho^2)^{1/2}, \quad \rho < a. \quad (19.43)$$

From this expression, we can derive the stress outside the fault ($\rho > a$), since $\sigma(\rho) = \mu \partial u / \partial \rho$:

$$\sigma(\rho) = \frac{\sigma_0 \rho}{(\rho^2 - a^2)^{1/2}}, \quad \rho > a. \quad (19.44)$$

Equations (19.43) and (19.44) describe the static distribution of slip inside the fault and stress outside (Fig. 19.11). Since the drop in stress is total, $\Delta \sigma_s = \sigma_0$, equation (19.43) establishes the relation between the slip and the stress drop. We must notice from (19.44) that a complete drop in stress inside the fault implies that the stress becomes infinite outside at its border ($\rho = a$), which is not physically possible.

19.5.2 The dynamic problem

The dynamic problem requires solution of the fracture problem as a function of time. The fracture front propagates with a certain velocity and, as it advances, material becomes fractured. Behind the front, the stress becomes zero for a total drop in stress or has a residual value σ_f that depends on the friction. Let us consider a simple case with a plane rupture front unlimited in the x_2 direction that advances in the x_1 direction with a constant velocity v .

The relation between the direction of the slip in the fracture plane and its direction of propagation defines three modes of fracture (Fig. 19.12). In mode I, *tensional fracture*, Δu , is normal to the fracture plane and to the direction of propagation. As the fracture propagates, the two sides of the fault separate and therefore the drop in stress is always total. In mode II, *in-plane shear fracture*, Δu , is contained in the fracture plane and has the same direction as that of its propagation. In mode III, *antiplane shear fracture*, Δu , is contained in the fracture plane with its direction normal to that of its propagation. From the point of view of seismology, mode I has little application, since earthquakes are assumed to be produced by shear fracture. In mode II, along the x_1 axis (defined as the direction of

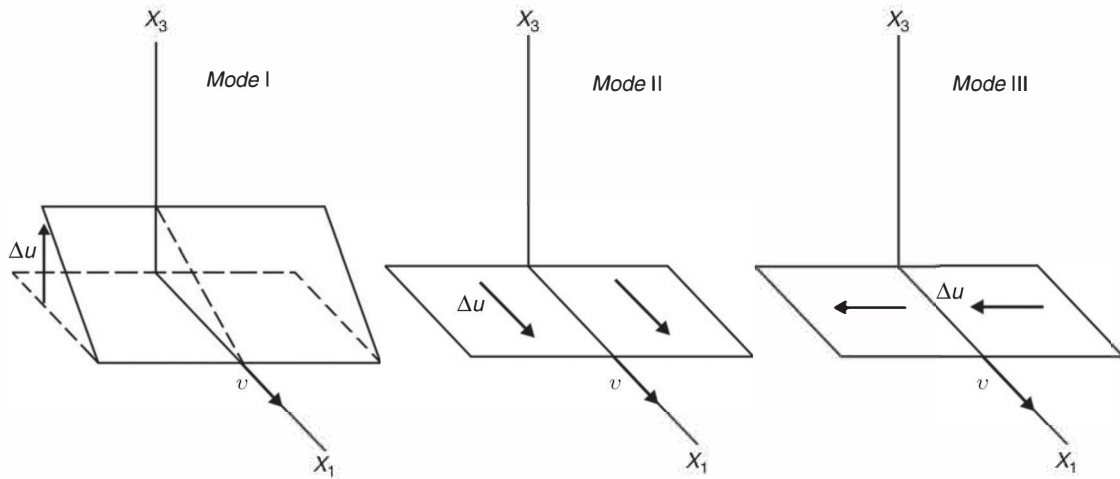


Fig. 19.12 Modes of fracture: mode I: tensional fracture; mode II: in-plane shear fracture; and mode III: antiplane shear fracture.

propagation of the fracture) we observe P and SV waves, whereas in mode III we observe only SH waves. The last case has a simpler solution.

The dynamic problem of the propagation of a fracture is centered on the energetic situation at the rupture front (Kostrov and Das, 1988; Fukuyama, 2009; Udías *et al.*, 2014). Let us consider a tensional fracture (mode I) whereby the stress drop is total. For the fracture front to advance, new fracture surface must be created and a certain amount of energy must be consumed. For this, an *elastic energy flux* from the part which has not been fractured to the fracture front is necessary. We call the energy necessary to create a unit of new fracture surface the *specific effective surface energy* or *Griffith energy* γ . The value of γ is a characteristic of each material. When the flux of energy to the rupture front G equals the energy necessary to create new fracture surface ($G = \gamma$), then fracture progresses. This is called the *Griffith fracture condition*. To produce an element of new fracture surface dS , the energy necessary is $2\gamma dS$ (the factor of two is due to there being two sides of a fracture). This energy comes from the stress drop $\Delta\sigma$ behind the rupture front and is given by $G = \Delta\sigma \Delta u dS$. For a perfect brittle fracture in elastic material, the material ahead of the fracture front (which is not yet fractured) is continuous ($u^+ = u^-$) and that behind (which has already been fractured) is discontinuous ($u^+ - u^- = \Delta u$). If the drop in stress is total, then, for each point behind the rupture front, the stress is zero ($\sigma = 0$). There is, then, a discontinuity at the rupture front and there is no transition between unfractured and fractured material.

The problem of the relation between the stress drop $\Delta\sigma$ and the slip Δu in the dynamic problem is not easy. In a simplified form, the problem consists of solution of the equation of motion (4.65) for the relative displacement (slip) of the two sides of the fault plane $\Delta u(x, t) = u^+ - u^-$, for $x < vt$, where x is the direction of propagation of the rupture with velocity v (for simple cases v is taken to be constant). The boundary conditions require

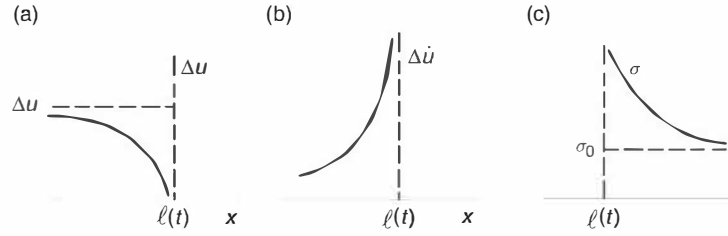


Fig. 19.13 The situation at the rupture front ($x = l(t)$) for dynamic fracture: (a) slip, (b) slip velocity, and (c) stress.

that, as the fracture front advances on the fault plane, the stress drops by $\Delta\sigma = \sigma_0 - \sigma_f$. If the fracture front is planar and propagates in the direction of x , with a constant velocity v , its position at each moment is given by $l(t) = vt$. The slip inside the fault can be written as

$$\Delta u(x, t) = \frac{\Delta u}{a} (v^2 t^2 - x^2)^{1/2}, \quad x < vt, \quad (19.45)$$

and the slip velocity is

$$\Delta \dot{u}(x, t) = \frac{V}{(v^2 t^2 - x^2)^{1/2}}, \quad x < vt, \quad (19.46)$$

where V is the *velocity intensity dynamic factor*. If the drop in stress is total ($\sigma_f = 0$), the stress inside the fault is zero and that outside, in a similar form to (19.44), is

$$\sigma(x, t) = \frac{K}{(x^2 - v^2 t^2)^{1/2}}, \quad x > vt, \quad (19.47)$$

where K is the *stress intensity dynamic factor* (not to be confused with the bulk modulus, Section 4.2). As the rupture front advances (vt increases in the x direction), the slip (19.45) increases from zero ahead of the rupture front to a constant value inside (Fig. 19.13a). The slip velocity (19.46) becomes infinite at the rupture front (Fig. 19.13b). The stress (19.47) becomes infinite when it approaches the rupture front from outside (Fig. 19.13c) and has a constant value or is zero inside. The dynamic factors K and V are related to the flux of energy from unfractured material to the rupture front

$$G = \frac{\pi}{2v} KV. \quad (19.48)$$

For mode III (antiplane), K and V are related by

$$K = V \frac{\mu}{2v} (1 - v^2/\beta^2)^{1/2}. \quad (19.49)$$

For a circular fracture that grows from its center, according to Madariaga (1976), the relation between Δu and $\Delta\sigma$ is

$$\Delta u(\rho, t) = \frac{\Delta\sigma}{\mu} C(v) v (t^2 - \rho^2/v^2)^{1/2}, \quad t > \rho/v, \quad (19.50)$$

where $v < \beta$, and $C(v)$ is a factor with a value near unity. The growth of the fracture is assured by the constant flux of energy from the material that has not yet been fractured to the fracture front. If the medium is homogeneous, then rupture, once it has started, cannot stop and grows indefinitely. This is due to the constant conditions of the material ahead of the rupture front.

Since, in an homogeneous medium, rupture, once it has started, cannot be stopped by itself, stopping must be introduced as a condition, for example, for a circular fault by imposing that the limit be at a predetermined value of the radius $\rho = a$. The beginning of the fracture must also be introduced as an added condition; for example, fracture starts when the applied stress exceeds a certain critical value. This value represents the maximum stress the material can support without breaking. Dynamic models of fracture allow determination of the slip and its propagation over the fault plane, from a specified drop in stress. Thus, we can solve for the elastic displacement field from the dynamic conditions in the source region. The solution of dynamic problems, even for simple cases, is difficult and in many cases they must be solved numerically.

19.6 Friction models of fracture

Recent dynamic models of fracture accord great importance to the problem of friction (Kostrov and Das, 1988; Cochard and Madariaga, 1994; Udías *et al.*, 2014). Thus an alternative model which can be used to represent the phenomenon of fracture is based on the friction between the two sides on a pre-existing fault surface. Before faulting occurs the two sides are kept locked by the action of *static friction stress* σ_s between them,

$$\sigma_s = \mu_s \sigma_n, \quad (19.51)$$

where μ_s is the *coefficient of static friction* and σ_n is the normal stress. When the applied stress overcomes the static friction ($\sigma > \sigma_s$), the two sides of the fault begin to move. Motion is resisted by the *kinematic friction stress* σ_k ,

$$\sigma_k = \mu_k \sigma_n, \quad (19.52)$$

where μ_k is the *coefficient of kinematic or dynamic friction*. In the simplest case, at time t_0 when the fault begins to move, the friction stress drops instantaneously from σ_s to σ_k and the slip between the two sides passes from zero to Δu . The effective or *dynamic stress drop* employed in producing the motion is

$$\Delta\sigma_d = \sigma_s - \sigma_k. \quad (19.53)$$

This value is different from the static stress drop $\Delta\sigma_s = \sigma_0 - \sigma_f$ of eq. (16.34). To overcome this, the friction stress acting on the fault has to increase from σ_0 (tectonic stress) to σ_s . The difference $\Delta\sigma_{ex} = \sigma_s - \sigma_0$ is called the *strength excess* or *fracture strength*.

In a simplified manner, the process in time of an earthquake can be shown as follows (Fig. 19.14): Acting tectonic stress σ_0 increases to σ_s at t_0 overcoming the friction stress and

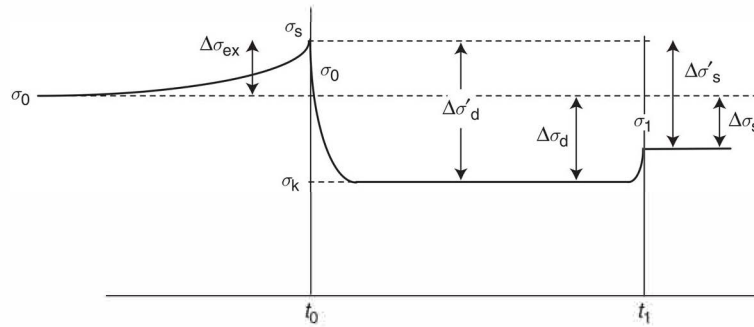


Fig. 19.14 Friction model of fracture. σ_0 tectonic stress, σ_s static stress before fracture, σ_f final friction stress that locked the fault; $\Delta\sigma_d$ and $\Delta\sigma_d'$ dynamic stress drop, and $\Delta\sigma_s$ and $\Delta\sigma_s'$ static stress drop.

then friction stress drops instantly to σ_k and the fault starts to move with slip Δu . At time t_1 friction stress rises to $\sigma_f > \sigma_k$ and motion stops. The final static stress drop $\Delta\sigma_s$ can be measured from σ_0 or from σ_s (Fig 19.14),

$$\Delta\sigma_s = \sigma_0 - \sigma_f, \quad (19.54)$$

$$\Delta\sigma_s' = \sigma_s - \sigma_f. \quad (19.55)$$

After the fault is locked the stress may continue to accumulate until it reaches again the value of σ_s and the fault will move in a new earthquake.

19.7 The complexity of a fracture

Under homogeneous conditions, the dynamic problem of the propagation of a rupture implies certain unrealistic border conditions concerning the stress and slip velocity and their nucleation and arrest. To solve some of these problems we must introduce inhomogeneities into the medium and complexities into the fracture process.

19.7.1 The cohesive zone

According to (19.47), there is a discontinuity of stress at the rupture front as it moves ($x = vt$). As we approach it from the outside, the stress becomes infinite and drops to zero or a constant value inside. Another discontinuity occurs for the slip velocity (19.46), which becomes infinite immediately behind the rupture front. This is a situation that is not physically possible, since no material of finite strength can sustain an infinite stress or move with an infinite velocity. These two inconsistencies of the homogeneous model follow from the fact that material is either purely elastic (unfractured) ahead of the front or fractured behind it. To avoid this situation we must consider the existence of a *transition zone* immediately ahead of the fracture front where material behaves in an inelastic way.

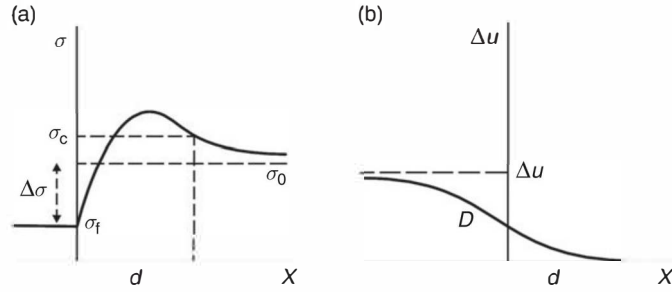


Fig. 19.15

The cohesive zone ahead of the rupture front: (a) cohesive stress and (b) the critical slip and length of the cohesive zone.

The first fracture model with a transition zone was proposed by Grigory I. Barenblatt (1959). The transition zone is called the *cohesive zone*. In this zone, cohesive forces act to oppose the advance of the fracture and hold the stress immediately ahead of the rupture front finite, eliminating the stress singularity. In the cohesive zone of width d , the stress has a finite mean value σ_c (the *cohesive stress*) that is larger than the applied tectonic stress σ_0 , and reduces to the friction stress behind the rupture front (Fig. 19.15a). The value of σ_c is related to the Griffith energy in the form,

$$\gamma = \frac{2\sigma_c^2 d}{\mu\pi C(v)}, \quad (19.56)$$

where $C(v)$ is a factor that depends on the velocity of fracture and has a value near unity for subsonic fractures. The slip does not become zero in the cohesive zone and there is no singularity in the slip velocity at the rupture front (Fig. 19.15b).

In the model with *weakening slip*, the cohesive stress is taken to be dependent on the slip (Ida, 1972; Palmer and Rice, 1973). This model assumes that the stress inside the fracture is a function of the slip $\sigma(\Delta u)$ in such a way that it has a finite value for $\Delta u = 0$, and decreases with increasing slip to a final value equal to the friction stress σ_f for Δu larger than a certain critical value $\Delta u = D$ (Fig. 19.16). The average stress, between $\Delta u = 0$ and $\Delta u = D$, equals the cohesive stress. The *critical value of the slip* D is related to the Griffith energy and the cohesive stress in the form

$$\gamma = \frac{1}{2} \sigma_c D. \quad (19.57)$$

This relation shows that the energy dissipated in the creation of a unit fracture surface equals the product of the cohesive stress and the critical slip. From (19.50) and (19.51), we can derive a relation between d and D :

$$d = \frac{\mu\pi C(v)}{4\sigma_c} D. \quad (19.58)$$

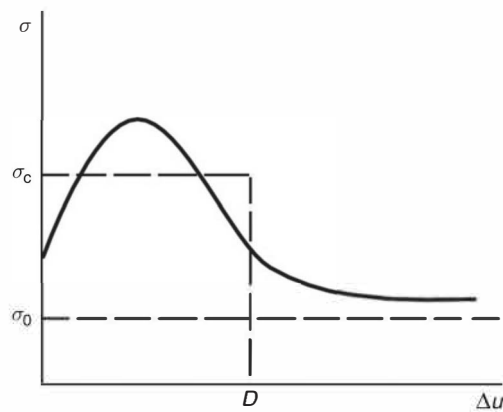


Fig. 19.16 Dependence of the stress on the slip in the slip-weakening model.

For earthquakes, values of d and D are small relative to the total dimensions of the fracture. For a fracture several kilometers long, d is only some meters and D some centimeters.

19.7.2 Barriers and asperities

We have seen that homogeneous models of fracture with uniform slip and constant rupture velocity are not very realistic. The simple fact that rupture must stop at the border of the fault indicates that the conditions cannot be totally homogeneous. In the Earth, faults cross rocks of various strengths, change direction often, and present jumps, joints, and bends (Scholz, 1990). Analysis of observed wave forms from earthquakes also reveals greater complexity than would be expected from homogeneous fractures. This is especially so for the complex form of high-frequency waves in the near field. Another item of evidence for the complexity of the source is the observation of practically constant values of stress drops (in the range 1–10 MPa) for earthquakes of magnitude larger than five. Since it has been shown in laboratory experiments that rocks can support larger stresses without breaking, the observed drops in stress are really average values for the whole of the fracture process. All these observations show that earthquake sources are *complex fracture processes*. Two models have been proposed to explain this complexity, namely, models with *barriers* and *asperities*.

The *barrier model* (Das and Aki, 1977; Aki, 1979) assumes that fracture takes place under uniform conditions of stress on the fault fracture, but with different strengths in the material (Fig. 19.17). Zones in the fault surface with high strength form *barriers* that make fracture propagation difficult or impede it. When the rupture front reaches a zone of barriers it stops and, if the barrier is sufficiently strong, it will not break.

Fracture may stop there or continue behind the barrier, leaving an unruptured zone on the fault plane. Once the fracture process has finished on the whole fault plane, stress is released in fractured zones and accumulated in the unbroken barriers. After the earthquake, the distribution of stress on the fault plane becomes heterogeneous. For a large earthquake there may be several barriers and the fracturing process is the sum of several ruptures separated

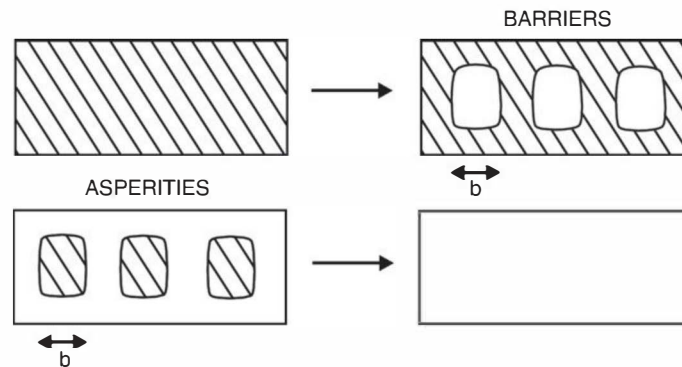


Fig. 19.17 Models of complex sources: barriers (top) and asperities (bottom).

by the barriers. Unbroken barriers may break later, giving rise to aftershocks (Section 20.2). Apostolos S. Papageorgiou and Aki (1983) have proposed a barrier model in which a rectangular fault is formed by several elementary circular fractures separated by barriers. In this model we can distinguish between the *global stress drop*, which is the mean value over the whole fault area, and the *local stress drop* due to the breakage of each individual elementary fracture. The latter is generally larger than the former. In this way we can explain the observed low values of drops in stress and the high energies in high-frequency waves of strong motion instruments in the near field.

The *model of asperities* (Kanamori and Stewart, 1978; Madariaga, 1979) consists of a fault with a heterogeneous distribution of stress on its surface, with zones of high and low values (Fig. 19.17). Zones with high stress are called *asperities*. This model takes into account the previous history of the accumulation of stress in certain zones of the fault (asperities) and the release of stress in other weaker zones. This process is achieved by the production of small earthquakes (foreshocks) that release stress in weak zones and accumulate it in strong ones. The breaking of strong zones or asperities where high stresses have accumulated constitutes the occurrence of the main large earthquake. The complexity of the source in this model is given by the fracture of several asperities in the main event. This corresponds to the complex STF in kinematic models. After the fracture of all asperities, the fault plane remains with a homogeneous distribution of residual frictional stress. This model explains the occurrence of foreshocks, but not aftershocks.

In both models, earthquakes are produced by complex fracture processes consisting of the breaking of several asperities or zones between barriers. Since neither model can explain both foreshocks and aftershocks, mixed models in which both asperities and barriers are present must be considered (Kostrov and Das, 1988). Barriers are zones that remain unbroken after the main earthquake and asperities are those that break with a high drop in stress. The distribution of stress on the fault plane is heterogeneous before and after an earthquake. Thus we can have both foreshocks that break the weak zones before the main shock and aftershocks that break the barriers that had been left unbroken. The complexity of the source is due both to the heterogeneous distribution of stress that is concentrated on the asperities and to the varying strengths of the barriers.

An important factor in the heterogeneity of the fault surface is the distribution of friction (Section 19.6). High and low values of friction along the fault plane contribute to the accumulation of stress. In the asperity model, motion starts when the applied stress overcomes friction and stops when the friction is larger than the applied stress. For constant friction, motion can only stop arbitrarily at the asperity's border. Authors of models introduce a weakening of the friction with velocity, that is, a decrease in friction with the slip velocity. In this situation the arresting of motion at the fault can be obtained from a *healing pulse* (Heaton's pulse) generated by friction itself (Heaton, 1990). Thus the rupture front advances by stress overcoming friction and is healed by friction behind the front. Energy is radiated, as the rupture front advances, from a narrow strip of the fault, while the rest of the fault has already been healed or has not yet broken. Thus, the motion does not need to be stopped by a healing pulse from the fault border, as we saw in kinematic models.

The search for more realistic models of earthquake sources leads to the consideration of complexities in the fracturing process, heterogeneities in the distribution of stress, strength, and friction along the fault surface, and the existence of a transition zone at the rupture front. These considerations increase the number of parameters necessary in order to define source models, by introducing the dimensions of asperities, the distance between barriers, length of cohesive zone, value of critical slip, the distribution of friction, etc. Also we must consider geometric irregularities of the fault surface such as branching, stepping, bending, and junctions that depart from the simple planar model (Andrews, 1989).

19.7.3 Acceleration spectra

We have seen how the dimensions of the source affect the form of the spectra of seismic waves in the far field. Complexities of the source influence the radiation at high frequencies and their effect can be observed in the accelerations in the near field, since they attenuate rapidly with distance. In complex models, rupture is not uniform, but rather has accelerations and decelerations, stopping and restarting with successive breaking of several elementary units (asperities). These irregularities result in the complexities observed in the accelerations of the near field at high frequency.

The spectrum of acceleration is related to that of displacement by a factor ω^2 . Thus, its form depends on ω^2 for low frequencies, corresponding to the flat part of the displacement spectrum, and is flat for frequencies higher than the corner frequency ω_c . For higher frequencies, we find a maximum value $\omega_{\max}$ or $f_{\max}$, approximately in the range 8–10 Hz, from which the spectrum decreases rapidly (Hanks, 1982) (Fig. 19.18a). This frequency has been related to the attenuation due to propagation in the upper part of the crust and to processes at the source. In the latter interpretation, this frequency is related to the smallest elementary dimension of asperities approximately about 200 m. Aki (1988), however, related $f_{\max}$ to the dimensions of the cohesive zone and the critical value of slip in the form

$$f_{\max} \simeq \frac{v}{d} \simeq \frac{4\sigma_c v}{\mu\pi C(v)D}. \quad (19.59)$$

The cohesive zone, with dimensions of about 100 m to 1 km, acts as a low-pass filter that is responsible for the attenuation of high frequencies.

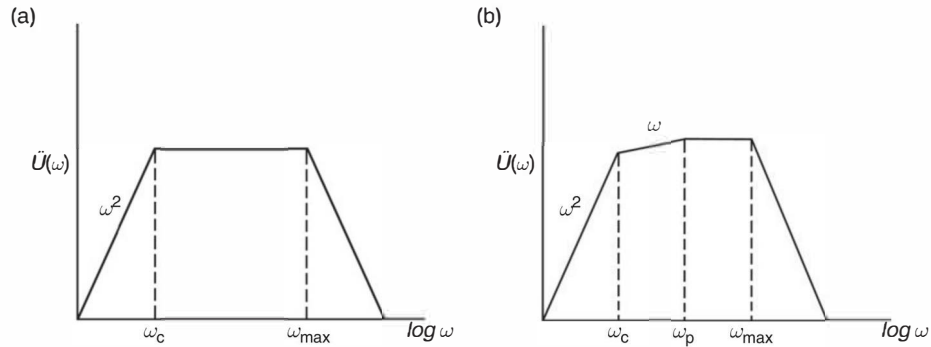


Fig. 19.18

The amplitude spectrum of acceleration in the near field (a) for a simple source and (b) for a complex source.

Other authors, such as Madariaga (1989), introduce another frequency ω_p with a change in the slope of the envelope of the spectrum which is related to the length b (or its average value $\bar{b}$) of elementary fractures (patches) that form the total fracture ($f_p \approx v/b$). This frequency has a value between the corner frequency ω_c and the maximum ω_{max} , which is due to attenuation effects (Fig. 19.18b). For small earthquakes, with only one elementary fracture, $\omega_p \approx \omega_c$ whereas for larger events $\omega_p > \omega_c$. The value of ω_c varies with the size of earthquakes, ω_p varies very little, and ω_{max} is practically constant. However, we must realize that the estimation of these frequencies from acceleration spectra is very much affected by attenuation. The study of source complexities has been made possible by digital strong motion records in the near field.

19.8 Summary

Realistic models of the source of an earthquake must consider the process of fracture on a finite surface, the slip produced on the fault, and the stress drop which has produced it. A simple kinematic model is Haskell's model with slip propagating at constant velocity in one direction of a rectangular fault. The amplitude spectrum of the elastic displacement field depends on the length of the fault, the velocity of fracture propagation, and the rise time. It is flat for low frequencies up to the corner frequency ω_c and decays for higher frequencies as $1/\omega^2$. A commonly used model for earthquakes of moderate magnitude ($M < 6$) is Brune's model of circular fault in which a stress pulse is applied instantaneously on the fault surface of radius a . The radius of the fault can be determined from the corner frequency of the spectra of the radiated S waves. Nucleation and arrest of the fracture propagation have been introduced in some kinematic models. Arrest is produced by the arrival of a cessation pulse from the limits of the fault or by the fracture itself propagating as a pulse.

Dynamic models consider the slip of the fault as a consequence of the stress conditions and the strength of the material at the fault. They are based on the theory of generation and

propagation of fractures in a stressed medium. The problem is centered on the energy situation at the fracture front. Thus we can find explicit relations for the slip as a function of the stress drop for certain models, for example, that of a circular fault. Another approach is the friction model of fracture where motion on a fault is produced when applied stress overcomes the friction that held its two sides locked. The dynamic problem leads to the introduction of complex situations on the fault. One is the existence of the cohesive zone at the fracture front, another the presence of barriers and asperities. Barriers detain the fault and accumulate stress while asperities are zones of stress concentration on the fault that break in main shocks. Barriers are needed to generate aftershocks and asperities to generate foreshocks.

19.9 Problems

- P19.1 For an earthquake, the corner frequency of the P wave is 0.023 Hz and of the S wave 0.1 Hz. If they correspond in a Savage model to ω_1^P and ω_3^S , what are the dimensions of the rectangular fault? Using only the S wave, what is the radius of the circular fault according to Brune's model? Compare both results. Use P wave velocity 6 km/s and Poisson's ratio 0.25.
- P19.2 The corner frequency of an S wave of velocity 3.4 km/s is 0.21 Hz. Determine the area corresponding to a square fault according to the Savage model and of a circular fault according to Brune's model. Do they agree?
- P19.3 From the first maximum of the directivity functions of the Rayleigh waves of two pairs of stations A and B (EMP_19.3), calculate the length of the fault and the rupture velocity, knowing that the phase velocity is 3.9 km/s and the angles between the azimuth of the stations and the strike of the fault are 60° and 160° .

Determination of the source mechanism for earthquakes consists of finding the parameters of the model used in its representation using observations of the radiated seismic waves. In [Chapter 16](#) we have seen the isotropic point source as the simplest point source model, in which the source is considered as a point from which waves propagate with equal amplitude in all directions. This model is the basis for the localization of the source in space and time. In the same chapter we have seen quantification of the size of an earthquake in terms of the magnitude, seismic moment, and energy.

In previous chapters we have seen several models representing the source mechanism, in particular, the point shear fracture or DC ([Chapter 17](#)), the first-order moment tensor ([Chapter 18](#)), and the extended models of fracture, kinematic and dynamic ([Chapter 19](#)). We will see in this chapter some of the methods used to obtain the parameters that define these models from the analysis of observations of seismic waves. We consider only the methods which use kinematic models.

20.1 Parameters and observations

Generally, in determination of focal mechanisms, parameters for the localization of the source (ϕ_0, λ_0, h , and t_0) are assumed to be known, although in some methods some (the focal depth) or all (centroid moment tensor) are determined anew. The size given by the magnitude (M_L, m_b, M_s or M_w) or the seismic moment (M_0) is independently determined, but, in many methods, M_0 is evaluated as part of the mechanism determination.

The simplest models of source mechanism are those with point sources. For a point shear dislocation, the orientation of the mechanism (strike, dip, and slip of the fault) is given by the angles ϕ , δ , and λ ([Section 17.1](#)). For the equivalent double couple force model, the orientation is given by that of the X and Y axes, that is, by the angles Θ_x , Φ_x , and Φ_y , or by the P and T axes, angles Θ_P , Φ_P , and Φ_T ([Section 17.2.2](#)). For the moment tensor representation (M_{ij}), if the source is not a shear fracture, for a deviatoric source it can be represented by five components of the moment tensor M_{ij} or six if there are changes in volume ([Chapter 18](#)). Including its size, a shear fracture point source or DC source is given by four parameters (M_0, ϕ, δ , and λ) plus four for its location, that is, a total of eight parameters. If the source is not a DC we need one or two more. If there is a rise time τ in the STF we have to add a new parameter and the total number is hence nine, ten, or eleven.

For an extended source we have to introduce the dimensions of its length and width (L, W) for rectangular faults or its radius (a) for circular faults. If there is propagation of the

fracture we introduce its rupture velocity (v) (Chapter 19). In kinematic fracture models, the value of the slip (Δu) can be obtained directly or from the seismic moment and dimensions. For a rectangular fault with a constant slip, rupture velocity, and rise time, there are five additional parameters (L , W , Δu , v , and τ), which with three for the source orientation and four for its location sum up to twelve parameters. Thus, even relatively simple models of the source require a considerable number of parameters for their definition. Complex models need new parameters such as the distribution of slip, rise time, and rupture velocity on the fault plane and on time, the number and dimensions of asperities, friction distribution, etc.

Observations used in the determination of focal mechanisms are the displacements of seismic waves recorded at several points on the Earth. Ground displacements are obtained from seismograms in graphical or digital form, corrected for the instrument's response (Chapter 3). Usually, methods are separated into those that use body and surface waves. Since waves propagate from the focal region to observation points, the elastic and anelastic properties of the Earth must be known *a priori*. Owing to the heterogeneous nature of the Earth, incomplete knowledge of its structure imposes certain limitations upon the investigation of the source. There is always a latent ambiguity about which characteristics of seismic waves are due to the source and which are effects of the propagating medium. There are ways to isolate these two effects, but there is a trade-off between details of the source that are to be determined and those of the medium that are supposed to be known. Simple models of the source, with observations at relatively large distances and low frequencies, are, in general, very little affected by propagation effects. However, very detailed models determined by observations from near distances and at high frequencies are more influenced by heterogeneities of the crustal structure.

An important factor to be considered in observations is that of the development of seismologic instrumentation (Chapter 3). The oldest seismographs (up to 1930) were mechanical, with smoked paper recording and having very low magnification (usually less than 1000) and not very precise time control. From 1930 onward, electromagnetic seismographs with photographic recording increased the amplification and were an important improvement. Installation of the WWSSN global network of seismographs in 1962 provided very good homogeneous observations for studies of mechanisms from that date up to 1990. Since 1990, modern digital high dynamic range broad-band instruments have been distributed globally and provide excellent data. At present, data can be obtained via the Internet from Data Centers (Section 3.6.1).

There are many methods for determination of focal mechanisms based on observations of body (P and S) and surface waves (Kasahara, 1981; Udías, 1991; Udías *et al.*, 2014). In the following sections, we will present the fundamentals of four of these, from the most simple to more complex, which are based on the signs of P wave first motions, wave-form analysis, moment tensor inversion, seismic wave spectra, and the distribution of slip on the fault plane.

20.2 P wave first motion polarities. Fault plane solutions

The first method developed to determine the focal mechanism is based on the sign or polarity of the first motion of P waves. The first authors to study data of this type were

Omori and Galitzin in around 1905, and in 1917 Toshi Shida was the first to recognize the alternating distribution of polarities in the four quadrants. Although there were several attempts by European and Japanese seismologists to use these data to study the source, the first operational method was proposed by Byerly (1928). The method is known as the *fault plane solution* and also as *Byerly's method*. The method was simplified by introduction of the *focal sphere*, proposed by Leonard P. G. Koning in 1942 and developed by Hirokichi Honda and Anne Reinier Ritsema among others in the 1950s. The focal sphere allows a simple graphical resolution of the problem using stereographic projections such as *Wulff* and *Schmidt nets* (named after Yuri V. Wulff and Walter Schmidt). Between 1950 and 1970, seismologists in the USSR used graphical methods extensively, adding to P polarities those of SV and SH waves (Balakina *et al.*, 1961). Because of the simplicity of polarity data, fault plane solutions are still very widely used.

The method is based on quadrant distribution with alternating signs (*compressions* and *dilations*) separated by two orthogonal planes of polarity of the radiation pattern of P waves produced by a shear fracture or double couple source (Section 17.6). The signs of the first motion of P waves are usually read from vertical component seismograms (upward for compressions, downward for dilations). The method consists of separating the four quadrants of compressions and dilations by two orthogonal planes, one of which is the fault plane. Owing to the seismic velocity dependence on depth, rays are curved and quadrants cannot be separated directly by using the distribution of observations on the Earth's surface. To correct for this effect, Byerly proposed extended distances, locating the points on the intersection of straight rays with the surface of the Earth. Today the concept of the *focal sphere* is more commonly used. The focal sphere is a sphere of homogeneous material and unit radius around the focus. Observation points are located at the surface of the focal sphere, corresponding to the rays that arrive at each station, by tracing their trajectories back to the focus (*ray back-tracing*). If an observing station is located at a distance Δ and azimuth ϕ' , at the surface of the focal sphere, its location is given by its azimuth ϕ' and take-off angle i_h measured from the vertical downward (Fig. 20.1) (Section 17.6).

To determine the take-off angle i_h which corresponds to a given distance Δ and focal depth h , it is necessary to know the structure of the medium of propagation. For teleseismic distances ($\Delta > 1000$ km), i_h can be calculated from the travel time curves $t(\Delta)$ for a focal depth h , by using Snell's law (10.5):

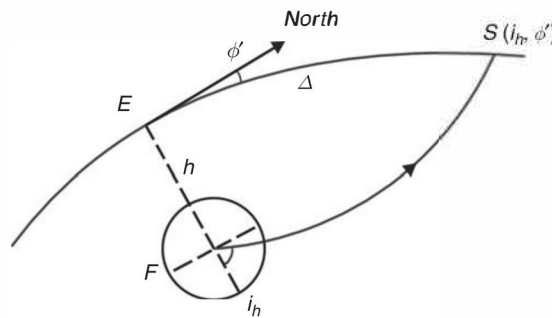


Fig. 20.1

The focal sphere and ray trajectory from the focus to a station.

$$\sin i_h = \frac{v_h}{r_h} \frac{dt}{d\Delta}. \quad (20.1)$$

It is necessary to know the velocity of P waves v_h at the focal depth h ($r_h = R - h$). Errors in this velocity result in errors in i_h and consequently in the mislocation of points on the focal sphere. For local or regional distances ($\Delta < 1000$ km), determination of i_h requires knowledge of the velocity distribution in the crust and upper mantle for the region. In general, layered models of constant velocity or with linear velocity distributions are used. However, models with velocity gradients are more convenient since they give a continuous relation of i_h to Δ , whereas layers with constant velocities give a discontinuous relation. Once we have determined values of ϕ' and i_h for each observation point, they are located on the surface of the focal sphere and the problem is reduced to that of a homogeneous medium. A reliable solution requires a sufficient number of observations distributed over azimuths and take-off angles so that the focal sphere is covered well.

The source model used corresponds to a point shear fracture or a double couple (Section 17.6). The orientation of the mechanism is given by that of the vectors $\mathbf{n}$ (the normal to the fault plane) and $\mathbf{l}$ (the direction of slip), or by that of the axes X and Y of the two couples of forces or also by P and T (the pressure and tension axes). Since all of these pairs are orthogonal unit vectors, the orientation is given by only three angles, $(\Phi_n, \Theta_n, \Phi_l)$, $(\Phi_X, \Theta_X, \Phi_Y)$, or $(\Phi_P, \Theta_P, \Phi_T)$. Also we can define the source in terms of the angles ϕ , δ , and λ for the orientation of the two planes, the fault plane (A) and auxiliary plane (B) (Fig. 20.2). The method cannot distinguish the fault plane from the

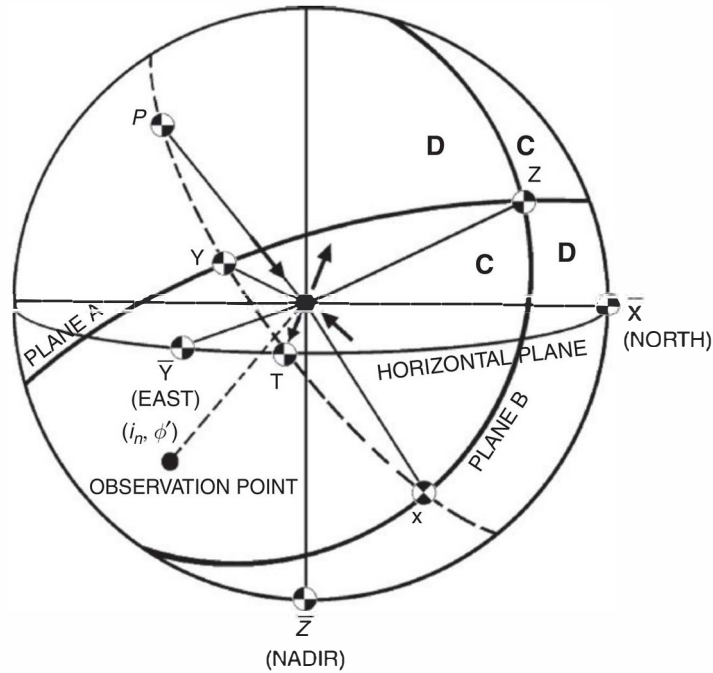


Fig. 20.2

The focal sphere with a vertical axis and a horizontal plane. Planes A and B separate quadrants of compressions and dilations. The locations of the X , Y , Z , P , and T axes are shown.

auxiliary plane due to the symmetry of the equations for the displacement field with regard to vectors $\mathbf{n}$ and $\mathbf{l}$ (Section 17.6).

20.2.1 Graphical methods

To solve the problem in graphical form, the focal sphere is projected onto a plane by means of stereographic projection; the most commonly used is the Schmidt or equal area projection (Fig. 20.3). Since, in most cases, observations correspond to rays leaving the focus downward, we project them onto the lower hemisphere. If there are very near stations or for very deep earthquakes at short distances, stations corresponding to upgoing rays are projected first onto the lower hemisphere along the diameter of the focal sphere. A point s in the lower hemisphere projects as s' on the plane AA' (the horizontal plane through the focus) (Fig. 20.3). The hemisphere projects as a circle of unit radius and a point on the sphere defined by (ϕ', i_h) projects as (ϕ', r) , where $r = \sqrt{2} \sin(i_h/2)$. Planes through the center project as great circles. In practice, the problem is solved by using a Schmidt net or its simulation on a computer screen.

The first step is the location of all observations (ϕ', i_h) on the projection at points (ϕ', r) , using different symbols for compressions and dilations (Fig. 20.4). Compressions and dilations are separated in the projection by two orthogonal planes into four quadrants alternating in sign. This is done by drawing first a plane (a great circle) and locating its pole (the normal through the center) and then drawing the second plane, such that it must pass through the pole of the first and its pole must be on the trace of the first plane (Fig. 20.4, where X is the pole of plane A and Y is that of plane B). Usually, several attempts are necessary in order to separate all compressions and dilations with a minimum number of inconsistencies. The quality of the solution is given by the score, which is the quotient of the number of correct observations and the total number of observations. Once we have obtained the solution, we measure the orientations of the two planes A and B in terms of angles ϕ , δ , and λ and those of axes X , Y , P , and T in terms of angles Θ and Φ (Fig. 20.4). We can see that all four axes are on the same plane (*plane of forces*), which is normal to the A and B planes.

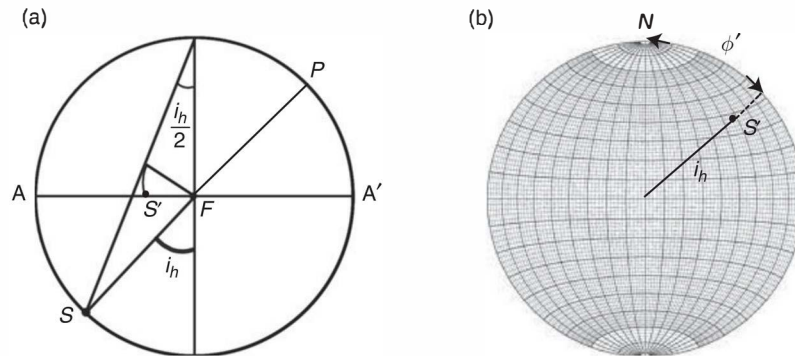


Fig. 20.3

Schmidt equal area projection (a) point S corresponding to a ray with take-off angle i_h at the lower hemisphere projects as S' on the horizontal plane. (b) Representation of the point $S'(\phi', i_h)$ on the Schmidt net diagram.

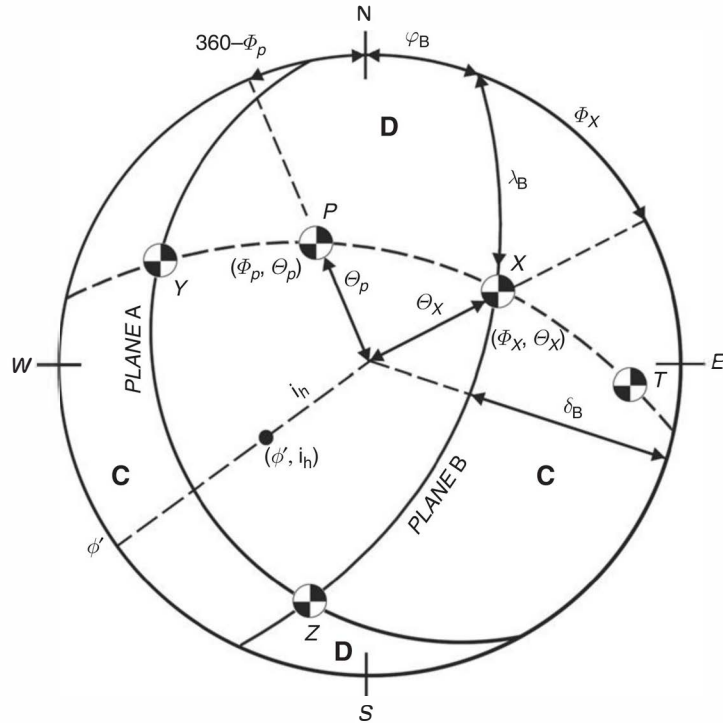


Fig. 20.4

Schmidt projection of the focal sphere. Planes A and B are the fault and auxiliary planes. The location of the X , Y , Z , P , and T axes and an observation point (ϕ', i_h) are shown.

The process we have described can be performed on the screen of a computer by using an interactive program. For example, a program implements the following steps (Bufo, 1994). Enter observations of the station name, azimuth (ϕ'), take-off angle (i_h), and P wave polarity; +1 for compression and -1 for dilation. Values of ϕ' and i_h have been found previously by another program. The program locates the points on the screen on an equal-area projection. A trial solution is given by specifying the orientation of X and Y , of P and T , or of ϕ , δ , and λ . The program draws the two orthogonal planes and determines the *score of the solution* (number of correct signs over total number). By successive steps, the solution is changed until the one with the maximum score is found.

20.2.2 Numerical methods

Since 1960, with the development of digital computers, several methods for the numerical determination of fault plane solutions have been proposed. The first formulation is that of Knopoff (1961), reformulated by Keichi Kasahara (1963), and applied in a computer program developed by Alan J. Wickens and John H. Hodgson (1967). This method seeks an orientation of the nodal planes which maximizes the probability of a correct observation at each station.

A computer program for fault plane solutions, widely used at present, is FPFIT (Reasenber and Oppenheimer, 1985). FPFIT computes double couple fault plane solutions

from P wave first motion data using a grid search method. The companion programs FPPLOT and FPPAGE plot the results on stereo nets for interactive viewing or for printing. There are additional programs in the package to create summary tables and to plot P and T axes for suites of mechanisms on stereo nets (<http://earthquake.usgs.gov/research/software/index.php>).

The program is based on an algorithm which uses the minimization of a one norm misfit function given, particularized for one earthquake, by

$$F^i = \frac{\sum_{k=1}^N \{ |p_o^k - p_t^{i,k}| w_o^k w_t^{i,k} \}}{\sum_{k=1}^N |w_o^k - w_t^{i,k}|}, \quad (20.2)$$

where N is the number of observations, p_o^k the observed polarity at station k , $p_t^{i,k}$ the theoretical polarity at station k corresponding to source model i , w_o^k observations weights estimated and assigned to the data, and $w_t^{i,k}$, the square root of the normalized theoretical P wave amplitudes at station k corresponding to model i (this downweights the observations near the nodal planes). The program compares the observed polarity at station k (p_o^k) with that calculated for a given orientation i of the source ($p_t^{i,k}$). The orientation of the source is given by the three angles ϕ , δ , λ , of one of the nodal planes. A perfect fit corresponds to $F^i = 0$ and a perfect misfit to $F^i = 1$. The program determines the variance of F^i and the 90% confidence limits and from these estimates the uncertainty of the parameters of the solution.

Another program, also available on <http://earthquake.usgs.gov/research/software/index.php> is HASH, which computes double couple earthquake focal mechanisms from P wave first motion polarity observations, and optionally S/P amplitude ratios. HASH is designed to produce stable high quality focal mechanisms, and tests solution sensitivity to possible errors in the first motion input and the computed take-off angles (Hardebeck and Shearer, 2002). There are other similar programs, for example FOCMEC, which uses a grid search of polarities and amplitude ratios S/P (www.iris.edu/pub/programs/focmec/ 2008 version).

We present briefly a method that uses maximization of a likelihood function and allows for individual and joint solutions for several earthquakes, giving estimations of the variance of the focal parameters (Brillinger *et al.*, 1980; Udías and Buforn, 1988). The method defines the probability that all observations are consistent with a determined orientation of the source as the product of the individual probabilities of each observation and looks for an orientation of the source that maximizes this total probability. The probability of observing a compression at station i is expressed as a function of the theoretical amplitude expected from an orientation of the source as

$$\pi_i = \text{prob}(Y_i = 1) = \gamma + (1 - 2\gamma)\Phi[A_i(\zeta_k)], \quad (20.3)$$

where $0 \leq \gamma \leq \frac{1}{2}$ represents reading errors and is given a small value ($\gamma = 0.01$) and $\Phi[A_i(\zeta_k)]$ is the Gaussian cumulative function of normalized theoretical amplitudes expected from a determined orientation of the source defined by its three parameters (ζ_1 , ζ_2 , and ζ_3) which correspond to ϕ , δ , and λ or Φ_P , Θ_P , and Φ_T . In this form, the probability is greater for values of A near unity (maximum amplitudes in the radiation pattern).

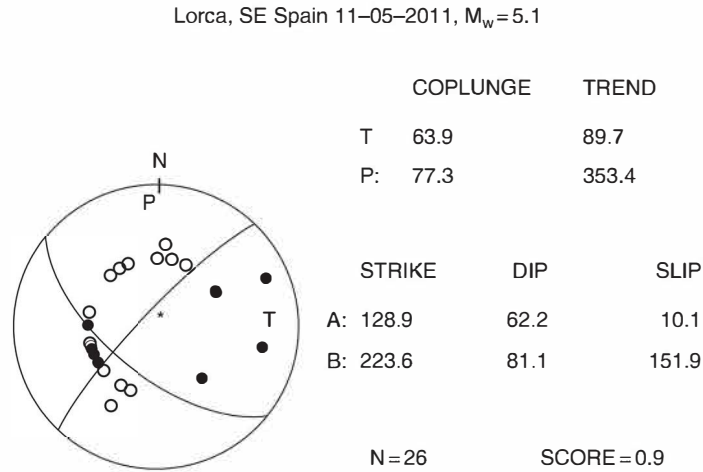


Fig. 20.5

The fault plane solution of the Lorca (SE Spain) earthquake, of 11 May 2011, $M_w = 5.1$.

If we have N observations Y_i for an earthquake, the probability that its mechanism corresponds to a determined orientation of the source given by ζ_k can be written as

$$P = \prod_{i=1}^N \left[\pi_i^{(1+Y_i)/2} (1 - \pi_i)^{(1-Y_i)/2} \right]. \quad (20.4)$$

The method searches for the parameters of the orientation of the model ζ_k that maximize the likelihood function $L = \log P$. Since L is a continuous function of ζ_k , its maximum is found by searching for the values of ζ_k that satisfy $\partial L / \partial \zeta_k = 0$. Once these values have been found, the method is used to calculate all the other parameters of the point source which are represented in the focal sphere (Fig. 20.5). The method determines the covariance matrix whose main diagonal gives the standard errors of the parameters and the information matrix. This method is generalized to find also joint solutions for groups of earthquakes with the same mechanism introducing a precision parameter ρ_j for each earthquake, which is useful in the study of mechanisms of aftershock sequences or swarms of earthquakes. The computer program MECSTA is available at: www.ucm.es/info/Geofis/g-sismolo/G-Sismolo.html.

20.3 Wave-form modeling

The method of P wave polarities uses the minimum information contained in seismic waves. Among the methods using more information are those based on the analysis of *wave forms* (Langston and Helmberger, 1975). These methods consist of comparison of P and S wave forms observed at various azimuths around the focus with those calculated from a point-source model. The comparison is made visually or by a method of minimization of the difference between observed and calculated waves (Nabelek, 1984). Wave-form analysis is frequently

applied to earthquakes with observations at teleseismic distances ($\Delta > 30^\circ$). Its application to local or regional earthquakes is more difficult because of the heterogeneity of the crustal structure. A solution for this problem is the use of *empirical Green's functions* (Hartzell, 1978). Empirical Green's functions are found from waves of small earthquakes located at the same place and recorded at the same stations as those that record the larger shock whose mechanism is being studied, and they are used to calculate the theoretical wave forms.

Wave-form modeling is based on the determination of theoretical or synthetic seismograms for P and S waves from a given model of the point source defined in terms of its depth, the orientation of the mechanism (DC), the source time function, and the seismic moment. For P waves of long periods, at teleseismic distances (30° to 90°) layering in the focal region and at the recording station does not affect wave forms. Thus the observed P waves are given by the sum of direct P waves and reflected pP and sP waves at the free surface above the focus. Because of the critical influence of the focal depth on these waves, this is a parameter that is evaluated as part of the analysis (Deschamps *et al.*, 1980; Okal, 1992).

According to (17.79), the displacement of P waves for an infinite medium is given by

$$u_j = \frac{M_0 f(t - r/\alpha)}{4\pi\rho\alpha^3 r} (n_k l_i + n_i l_k) \gamma_i \gamma_k \gamma_j. \quad (20.5)$$

For a shear fracture, the components of the vectors $\mathbf{l}$ and $\mathbf{n}$ can be expressed in terms of the fault angles ϕ , δ , and λ , according to (17.99)–(17.104). The STF $f(t)$ represents the time dependence of the slip velocity and γ_i are the direction cosines of the ray that arrives at the station and can be given in terms of its azimuth ϕ_s and take-off angle i_h at the focus (17.105)–(17.107).

If the focus is very deep, reflected pP and sP waves are separated from direct waves and we need to consider direct waves only (Fig. 20.6). The vertical component of a P wave at an observation point on the Earth's surface is given by

$$u_z^P = \frac{M_0 f(t - r/\alpha)}{4\pi\rho\alpha^3 r} R^P(\phi, \delta, \lambda, i_h) g(\Delta) G(r, Q) C_z(i_0), \quad (20.6)$$

where r is the distance along the ray from the focus to the station, $R^P(\phi, \delta, \lambda, i_h)$ is the radiation pattern for P waves, ϕ is now the azimuth at the focus measured from the trace of

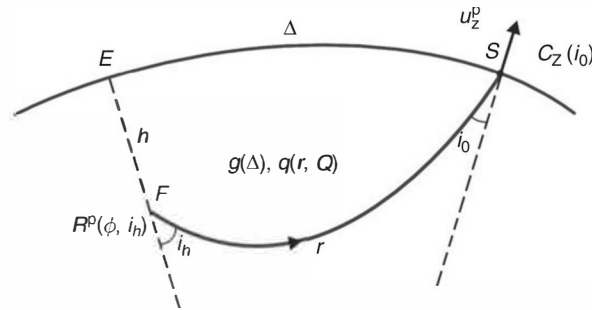


Fig. 20.6

Ray trajectories of P waves from the focus to a station and elements that are used in wave-form modeling.

the fault ($\phi = \phi_{\text{fault}} - \phi_{\text{station}}$), i_h is the take-off angle at the focus, i_0 is the angle of incidence of the ray at the station, $g(\Delta)$ is a factor due to geometric spreading of the wave front (Section 5.9), $G(r, Q) = \exp[-\omega r/(2\alpha Q)]$ is the anelastic attenuation (Section 15.3), and $C_Z(i_0)$ is the effect of the free surface on amplitudes (Section 6.5).

In terms of ϕ , δ , and λ , the pattern of the radiation of P waves is given by

$$R^P = A(3 \cos^2 i_h - 1) - B \sin(2i_h) - C \sin^2 i_h, \quad (20.7)$$

and that of SV waves is given by

$$R^S = -\frac{3}{2}A \sin(2i_h) - B \cos(2i_h) - \frac{1}{2}C \sin(2i_h), \quad (20.8)$$

with

$$A = \frac{1}{2} \sin \lambda \sin(2\delta), \quad (20.9)$$

$$B = \sin \lambda \cos(2\delta) \sin \phi + \cos \lambda \cos \delta \cos \phi, \quad (20.10)$$

$$C = \sin \delta \cos \lambda \sin(2\phi) - \frac{1}{2} \sin \lambda \sin(2\delta) \cos(2\phi). \quad (20.11)$$

The STF $f(t)$ has, for a simple source, a triangular or trapezoidal form, and, for a complex source, a combination of triangular and trapezoidal functions (Section 17.7). The duration of the source action is given by the rise time (τ) in the first case and by the sum of all rise times in the second.

In general, the displacement $u_Z^P(t)$ at the surface is the sum of P, pP, and sP waves which arrive near in time:

$$\begin{aligned} u_Z^P(t) = & \frac{M_0}{4\pi\rho_h\alpha_h^3r} g(\Delta) G(r, Q) C_Z(i_0) [R^P(\phi, i_h) f(t - t_P) \\ & + R^P(\phi, \pi - i_h) V_{pP} f(t - t_P - \Delta t_{pP}) \\ & + R^S(\phi, \pi - j_h) V_{sP} f(t - t_P - \Delta t_{sP})] \end{aligned} \quad (20.12)$$

The first term corresponds to the direct P wave, the second to the pP wave, and the third to the sP wave (Fig. 20.7). R^P and R^S are the normalized radiation patterns for P and SV waves (20.7) and (20.8). For pP and sP waves, rays take-off from the focus in the upward direction and correspond to angles of incidence $\pi - i_h$ and $\pi - j_h$, and, according to Snell's law, $\sin j_h = (\beta_h/\alpha_h) \sin i_h$. If t_P is the arrival time of P waves, then the delays with respect to this time of pP and sP wave arrival are calculated using the approximation that direct and reflected rays have the same take-off angles at the focus and the surface (Fig. 20.7a). Then, from F and F', the travel times of the two rays are the same and the delay $\Delta t_{pP} = (l_1 + l_2)/\alpha$, where $l_1 + l_2 = 2h \cos i_h$. The delays of sP waves are derived in a similar way:

$$\Delta t_{pP} = t_{pP} - t_P = \frac{2h \cos i_h}{\alpha_h}, \quad (20.13)$$

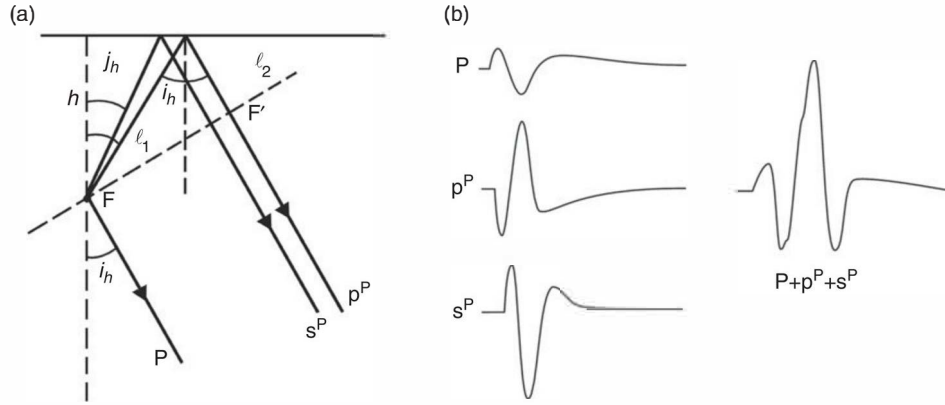


Fig. 20.7 Ray trajectories at the focus (a) and at a station (b) for P, pP, and sP waves.

$$\Delta t_{sP} = t_{sP} - t_P = h \left(\frac{\cos j_h}{\beta_h} + \frac{\cos i_h}{\alpha_h} \right). \quad (20.14)$$

V_{PP} and V_{sP} are the coefficients of reflection at the free surface for an incident P and reflected P wave, and an incident SV and reflected P wave. These coefficients are for the vertical component of displacement and can be derived from (6.100) and (6.107) (Section 6.4). The observed P wave is the sum of the three waves (Fig. 20.7b).

In the process of modeling wave forms, the parameters which are adjusted are the seismic moment M_0 , orientation of the source (ϕ , δ , and λ), source time function $f(t)$, and focal depth h . The rest of the parameters are assumed to be known and are kept constant except for those that vary with the focal depth.

The method consists of selecting observations of P waves at different azimuths within the accepted range of distances. At present the best results are obtained by using broad-band digital records. Observed wave forms must be transformed into ground displacements by correcting for the instrument's response (Section 3.6.3). Theoretical wave forms are calculated by assuming some initial values of the variable parameters. As preliminary values the orientation of the source from the fault-plane solution and focal depth from hypocentral determination may be used. Values of focal parameters are systematically changed to improve the fit between theoretical and observed wave forms. Sometimes the value of the anelastic coefficient Q is also changed. On increasing h , delays between arrivals of P, pP, and sP waves increase, and on varying Q the width of the pulses is changed. The fitting is done visually by comparing theoretical and observed wave forms or numerically by an iterative procedure in which the residuals of amplitudes are minimized by a least squares method. A method in which the parameters are adjusted by a least squares procedure is known as *wave-form inversion* (Nabelek, 1984; MacGaffrey *et al.*, 1991). An example of wave-form modeling is given in Fig. 20.8. The method is also applied to S waves by separating SV and SH components beforehand. Theoretical seismograms are calculated directly in the time domain or are first calculated in the frequency domain and then transformed to the time domain. A problem with these methods, especially if only

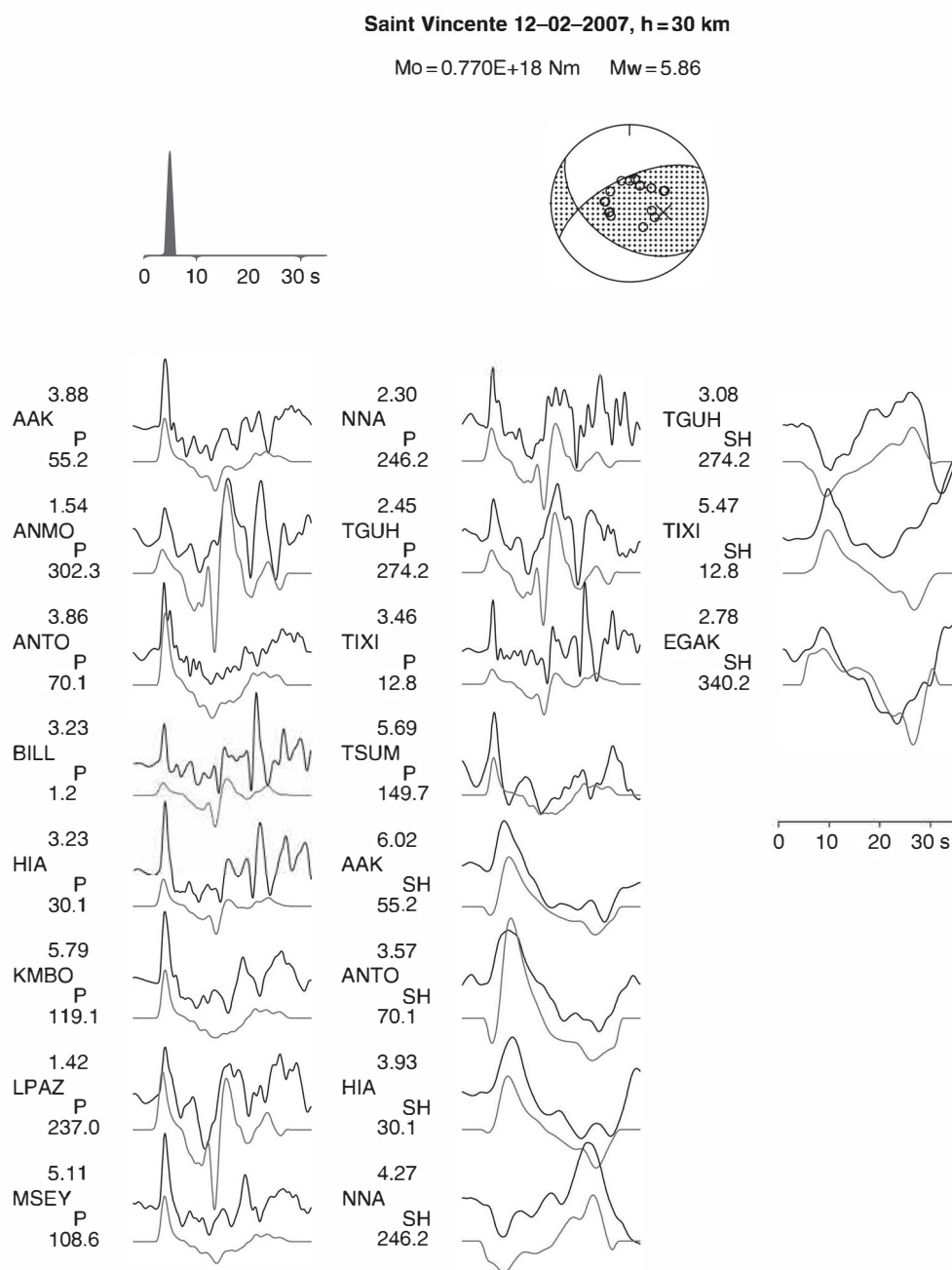


Fig. 20.8 Modeling of P and SH waves for the Saint Vincent earthquake of 12 February 2007. (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

a few stations with a poor coverage of azimuths are used, is how to distinguish assigning the observed complexity of wave forms to the source that may actually belong to the propagating medium.

20.4 Inversion of the moment tensor

In Section 18.7, we saw how formulation of the source mechanism in terms of the moment tensor allows a linear inversion of its components from observations of seismic wave displacements. The most common problem concerns inversion of the first-order time-independent moment tensor that corresponds to a point source. The first method used observation of free oscillations (Gilbert and Dziewonski, 1975). Methods for inversion of surface waves in the time and frequency domains (Mendiguren, 1977; Kanamori and Given, 1981) and of body waves have been proposed (Stump and Johnson, 1977; Strelitz, 1978). A generalized method applicable both to body and surface waves uses summation of normal modes and the concept of the centroid moment tensor (CMT) (Section 18.7) (Dziewonski *et al.*, 1981). This method was applied on a routine basis at Harvard University to sufficiently large earthquakes and at present is continued by the Global Centroid Moment Tensor Project (www.globalcmt.org) led by Goran Ekstrom. A method that also uses the CMT inversion with both body and surface waves is applied in a routine form by the U.S. Geological Survey to all earthquakes of magnitude larger than 5 (<http://neic.usgs.gov/equarchives/sopar>) (Sipkin, 1982). Most inversion methods of the moment tensor assume the source to be purely deviatoric and afterwards separate the DC part of the solution (Section 18.4).

For regional distances the inversion of the moment tensor may be carried out from body or surface waves (Cesca *et al.*, 2010; Hermann, 2013). Software for regional distances may be retrieved from websites: www.eas.slu.edu/eqc/eqcpcps.html; <http://crack.seismo.unr.edu/htdocs/students/Ichinose>; and <http://kinherd.org/kiwitools/>.

20.4.1 The inversion of Rayleigh waves

As an example of the methodology followed in moment tensor inversions, we present the basic procedure followed in the inversion of Rayleigh waves. Usually vertical-component seismograms recorded at several stations are used. The data must be reduced to ground motion by correcting for the instruments' responses. If all data are reduced to a common distance, in the frequency domain, then the complex spectra are functions only of the azimuth of each station ϕ measured from the trace of the fault and the frequency ω :

$$U(\omega, \phi) = A(\omega, \phi)e^{i\Phi(\omega, \phi)} = R(\omega, \phi) + iI(\omega, \phi). \quad (20.15)$$

According to (18.54), the spectra can be expressed as the product of the spectrum of the seismic moment and the derivatives of Green's function:

$$U_k(\omega, \phi) = M_{ij}(\omega)G_{ki,j}(\omega, \phi). \quad (20.16)$$

The real and imaginary parts of the spectra are linear combinations of the components of M_{ij} and those of $G_{ki,j}$. The latter are called the *excitation functions* G_k , which for a purely deviatoric source can be reduced to seven components. These functions are calculated for an appropriate model of a layered medium for $\phi = 0$, and for other values of ϕ are given by

their product with the sines and cosines of ϕ . The spectrum of the source time function can be separated into the form, $M_{ij}(\omega) = M_{ij}F(\omega)$, assuming that all of the components have the same time dependence (synchronous moment tensor). For low frequencies, the problem is practically independent from the frequency, since we are dealing with the flat part of the spectrum. In this case, the real and imaginary parts of the spectrum for the vertical components of displacements and a deviatoric tensor are given by

$$R = G_1[M_{xy} \sin(2\phi) - \frac{1}{2}(M_{yy} - M_{xx}) \sin(2\phi)] + G_2 M_{zz} - \frac{1}{2} G_1 M_{xx}, \quad (20.17)$$

$$I = G_3(M_{yz} \sin \phi + M_{zx} \cos \phi), \quad (20.18)$$

where the components of the moment tensor are referred to the geographic axes (Section 18.1). The variables are the six components of the moment tensor. For their solution, we need observations at N stations at different azimuths so that the problem is well conditioned. The resulting system of $2N$ equations with six unknowns is solved by a least squares procedure.

The problem can also be solved by using the spectral amplitudes of Rayleigh waves $A(\omega, \phi)$. In this case, their relation to the components of M_{ij} is not linear. If the six components of the moment tensor are M_i , $i = 1, \dots, 6$, then the spectral amplitudes can be expressed as $A(\omega, \phi, M_i)$. The problem is solved by taking a Taylor expansion about some initial known values M_i^0 :

$$A(\omega, \phi, M_i) = A(M_i^0) + \sum_{k=1}^6 \frac{\partial A}{\partial M_k} \delta M_k. \quad (20.19)$$

The unknowns are now the six values of the corrections δM_i , and the solution is given by $M_i = M_i^0 + \delta M_i$. Just like in the previous case, we need observations from N stations at different azimuths and the problem is solved by a least squares method. The process is repeated by an iterative method until it converges to a minimum error solution. The error is defined as the sum of the squares of residuals between observed and calculated spectral amplitudes at each iteration.

In these methods the moment tensor is assumed to be deviatoric and hypocenter coordinates and origin times are supposed to be known. Since the focal depth affects the excitation functions and its value is often not very precise, it is incorporated into the problem as another variable. In the CMT method, centroid hypocenter coordinates and the origin time are also incorporated into the problem, as four added unknowns that are recalculated together with the five components of the moment tensor. As already mentioned, coordinates obtained from the minimum error solution correspond not to the point of initiation of the fracture, as do those obtained from arrival times, but rather to the centroid of the source area. This relocation of the hypocenter results in a better estimation of the moment tensor's components.

In all methods of inversion of the moment tensor, an appropriate model of the Earth structure must be used to calculate the excitation functions. Errors in this model result in errors in the moment tensor obtained. For low frequencies this is not important, but it

becomes a problem for small distances and relatively high frequencies. The non-DC part of the solution is in many cases due to these errors, rather than to the source itself. As has already been mentioned, moment tensor inversions give the five components of a deviatoric tensor and must be separated into a DC part and a non-DC part (Section 18.4). From the DC part, the parameters of the orientation of the source are obtained according to (18.24). The method gives also the STF, interpreted now as the rate of release of the moment as a function of time. In general, solutions are more stable when observations are taken from many different azimuths.

20.5 Amplitude spectra of seismic waves

The amplitude spectra of body and surface waves can be used to obtain certain source parameters such as scalar seismic moments (M_0) and source dimensions (L or a) (Hanks and Wyss, 1972). According to Sections 19.2 and 19.3, the amplitude spectra of seismic waves depend on the source dimensions, practically independently from the model of the source. The seismic moment is determined from the value of the flat part of the spectrum for low frequencies of body and surface waves, corrected for the instrument's response, the radiation pattern, and all factors that affect wave propagation. For P waves, the seismic moment is given by

$$M_0 = \frac{4U_Z^P \pi \rho \alpha^3 r \exp[\omega r / (\alpha Q_\alpha)]}{g(\Delta) C_Z(i_0) R_P(\phi, \delta, \lambda, i_h)}, \quad (20.20)$$

where U_Z^P is the mean spectral amplitude for low frequencies for the vertical component of P waves, r is the distance along the ray from the focus to the station, Q_α is a quality factor for P waves, $g(\Delta)$ is the geometric spreading, $C_Z(i_0)$ is the effect of the free surface on amplitudes, and $R_P(\phi, \delta, \lambda, i_h)$ is the pattern of radiation corresponding to the orientation of the source given by ϕ , δ , and λ , where ϕ is the azimuth from the fault trace to the station and i_h is the take-off angle at the focus. The expression for S waves is similar. For Rayleigh waves the expression is

$$M_0 = \frac{U_Z^R (2\pi R)^{1/2} \exp(\gamma_R R)}{k_R^{1/2} N_Z(R, h) R_R(\phi, \delta, \lambda)}, \quad (20.21)$$

where U_Z^R is the mean spectral amplitude for low frequencies of the vertical components of Rayleigh waves, R is the distance from the epicenter, γ_R is an attenuation factor, k_R is the wave number, $N_Z(R, h)$ is the excitation function corresponding to the depth h , and $R_R(\phi, \delta, \lambda)$ is the radiation pattern. Expressions (20.20) and (20.21) can be simplified by making appropriate approximations for some factors.

As we saw in Sections 19.2 and 19.3, the source dimensions are related to the corner frequency ω_C (Fig. 20.9a). The widths of the time pulse and of the spectrum are related by $\Delta t \approx 1/\omega_C$ (Fig. 20.9b). If the observed spectrum is drawn on a logarithmic scale it can be fitted by two straight lines, one parallel to the frequency axis for low frequencies and another of slope -2 for high frequencies. The intersection of these two lines marks the

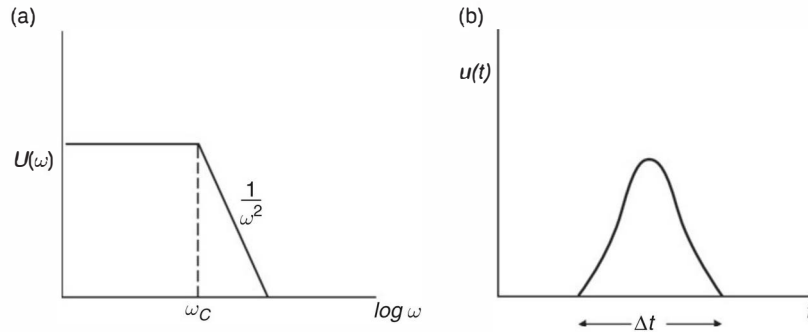


Fig. 20.9

Effects of source dimensions on the spectrum (a) and the pulse width (b) of P waves.

corner frequency. For a rectangular fault of length L and width W , according to the Savage model (19.24),

$$(LW)^{1/2} = \frac{1.7\alpha}{\omega_c^p} = \frac{3.8\beta}{\omega_c^s}. \quad (20.22)$$

For a circular fault, the radius, according to Brune's model (19.31),

$$a = \frac{2.33\beta}{\omega_c^s}. \quad (20.23)$$

In this form, both the seismic moment and the source dimension can be estimated from amplitude spectra (Fig. 20.10). Estimations of the seismic moment are very stable, if we use wave lengths larger than the source dimensions. Estimates of dimensions for small earthquakes can be affected by attenuation. For this reason, determinations must be made from data of several stations and a mean value taken. Once the seismic moment and dimensions are known, we can estimate values of the apparent average stress and stress drop (Section 16.6).

20.6 Determination of the slip distribution over the fault plane

Once the orientation and dimensions of the fault plane are known (ϕ , δ , λ , L , and W) we can determine the distribution of slip rate $\Delta \dot{u}$ over the fault plane from the analysis of wave forms. There are several methods to solve this problem (for example, Kikuchi and Kanamori, 1982, 1991; Hartzell and Heaton, 1983; Archuleta, 1984; and Fukuyama and Irkura, 1986). As we saw in Section 19.5, we can first consider the change in the amplitude of the slip rate with constant slip direction. The rectangular fault plane is discretized by N points in the length direction and M in the width, using increments of Δx and Δy length, so that $L = N \Delta x$ and $W = M \Delta y$. For constant rupture propagation velocity v in the x direction we can write (19.1) as

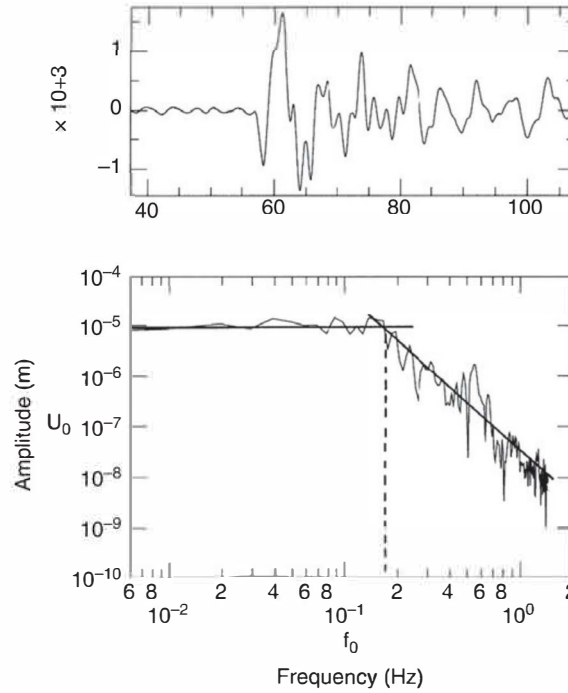


Fig. 20.10

Determination of the seismic moment and the radius of the source from the amplitude spectrum of P waves for the Azores earthquake of 27 June 1997; $M_0 = 3.5 \times 10^{17}$ Nm, $r = 10$ km.

$$u^p(i_h, \phi, r_0, t) = \frac{R_p(\phi, \delta, \lambda, i_h, \phi') \mu}{4\pi\rho\alpha^3 r_0} \sum_{i=1}^N \sum_{j=1}^M \Delta \dot{u}_{ij} \left(t - \frac{r_0}{\alpha} + \frac{i\Delta x}{v} + \frac{i\Delta x \sin i_h \cos \phi'}{\alpha} \right). \quad (20.24)$$

In this equation we have put the initial point of the rupture propagation ($i = 1, j = 1$) at the upper left corner of the fault plane.

For a general case, the initial point can be located anywhere on the fault plane, usually, at the hypocenter. If rupture expands from this point in the x and y directions with the same constant velocity v , the slip rate can now be written in the form,

$$\Delta \dot{u}_{ij} \left[t - \frac{r_0}{\alpha} - \left(\frac{i\Delta x}{\alpha} + \frac{j\Delta y}{\alpha} \right) \left(\frac{\alpha}{v} - \sin i_h \cos \phi' \right) \right]. \quad (20.25)$$

To ensure that the motion on the fault stops at its border, a boundary condition must be added that $\Delta \dot{u}$ diminishes to zero as it approaches the limits of the fault. The locations, where higher values of $\Delta \dot{u}$ are found, correspond to those of the presence of asperities in the inhomogeneous dynamic model (Section 19.7.2).

For a shear fracture, if the slip direction changes over a fault plane with fixed orientation, we let the rake angle λ vary together with the slip amplitude. For rupture propagation only in the x direction, according to (19.1), we can write

$$u^p(i_h, \phi, r_0, t) = \frac{\mu}{4\pi\rho\alpha^3 r_0} \sum_{i=1}^N \sum_{j=1}^M R_p(\phi, \delta, \lambda_{ij}, i_h, \phi') \Delta \dot{u}_{ij} \left(t - \frac{r_0}{\alpha} + \frac{i\Delta x}{v} + \frac{i\Delta x \sin i_h \cos \phi'}{\alpha} \right). \quad (20.26)$$

The standard procedure to model extended sources is to assume a finite rectangular fault plane divided into a set of contiguous sub-faults of area A_j , slip rate moment $\Delta \dot{u}$, and source-time function $f_j(t)$.

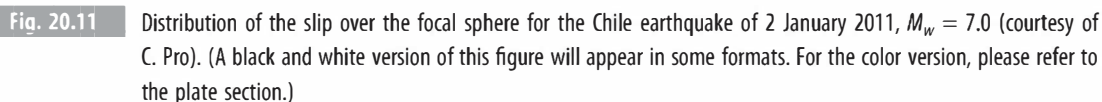
As we saw for the case of the point source, the seismic wave radiated at the far field for shallow far field sources has to include not just the direct P wave emitted by the source but also the pP and sP waves reflected by the free surface above the source (Section 20.3). For extended sources, for each of the phases P, pP, and sP, the observed wave form at the i th station is given by the sum of the contributions from each sub-fault. To ensure that the motion on the fault stops at its borders, as was mentioned, a boundary condition must be added forcing $\Delta \dot{u}$ to diminish to zero as it approaches the limits of the fault.

In some methods, we can also allow the rake of the slip direction to change over the fault plane of fixed orientation. Thus, we let the rake angle λ vary together with the slip-rate amplitude $\Delta \dot{u}$. To avoid inconsistencies in the direction of slip over the fault surface, λ is only allowed to change over a certain range around a given initial value. In theory, we could retrieve all the parameters involved in the slip distribution over the fault plane from the inversion of body waves, but, in practice, some parameters are fixed and the inversion is only carried out for a few of them. For example, information about the dimensions of the fault must be assumed from other considerations. Normally, in a first step we obtain the fault-plane orientation and depth using a point source (Section 19.3). Fixing these values we do the inversion for the distribution of slip rate on an extended source.

The amount of slip and the time history of slip are determined at each grid point, using a fixed rupture velocity and from these values we obtain the synthetic seismograms. The solution is found by minimizing the fit between synthetic and observed data. Normally, different sized grid, rupture velocities, time of initiation or onset of the rupture, etc. are tested. It is important to take into account that the grid size and rise time are related to the rupture velocity. One of the most robust methods for this type of inversion is the algorithm developed by Masayuki Kikuchi and Kanamori for body waves at teleseismic distances (30°–90°) (www.eri.u6tokyo.ac.jp/ETAL/KIKUCHI). An example of the application of this method is given in Fig. 20.11.

20.7 Summary

The source mechanism of earthquakes is defined in terms of the parameters which define the motion on the fault and can be found from the analysis of seismic waves. There are a variety of methods proposed to solve this problem. Several methods have been presented for kinematic models. The simplest method is that based on the distribution of the polarities



of the P waves and gives the orientation of the fault plane (fault plane method) using a shear fracture point source. This is a widely used procedure solved by graphical or numerical procedures. Other common methods for solving the problem are those that use analysis of the wave form of body waves.

The components of the moment tensor can be obtained from the inversion of body or surface waves in the time or frequency domain. Usually a deviatoric source is assumed (five components of the moment tensor) which is afterwards separated into DC and CLVD source components. When the CMT is used the methods provide the location and origin time of the centroid.

Dimensions of the source (length of rectangular fault or radius of circular fault) can be obtained from the amplitude spectrum of body or surface waves. Finally the distribution of the slip or slip rate and its orientation on the fault surface can be obtained from the analysis of wave forms.

20.8 Problems

P20.1 The earthquake of 19 June 1982, $M = 7$, located at 74.2 W, 16.2 N (offshore from El Salvador), has been recorded at teleseismic distances by 14 stations. Their azimuths, take-off angles, and polarities of P waves (compressions and dilations) are

ALQ	327	43	C	GUMO	294	20	D	SLR	112	20	D
DBF	124	35	D	KEV	18	22	C	SNZO	230	20	D
BOCO	118	57	C	LON	327	38	D	TATO	322	20	D
BER	30	24	C	NWAO	227	7	D	TOL	52	25	C
GRFO	40	22	C	SCP	18	43	C				

Using the equal area projection (Schmidt net, EMP_20_1) or a computer program (websites given in the text):

- situate the observations on the focal sphere;
- find the two nodal planes and give their orientations in terms of ϕ , δ , and λ ; and
- find the angles Φ and Θ of the axes X , Y , Z , P , and T .

P20.2 An earthquake at Granada (Spain) on 20 June 1994, $M = 4.5$, has been recorded at regional distances by 21 stations. Their azimuths, take-off angles, and polarities (compressions and dilations) are

FOR	305	71	C	CAC	261	36	D	ALOJ	135	51	C
TOR	37	42	D	SAL	177	51	C	ASMO	35	80	D
ALM	69	20	D	ELOJ	136	67	C	ACHM	247	70	D
ESC	272	82	D	ERON	277	71	D	RESI	320	78	C
ROM	157	46	C	ECOG	57	77	D	PARA	197	89	C
FRA	299	52	C	APHE	287	88	D	MON	78	78	D
SIL	88	86	D	ATEJ	358	60	C	DOL	28	88	D

Using the equal area projection (Schmidt net, EM20.1) or a computer program (websites given in the text):

- situate the observations on the focal sphere;
- find the two nodal planes and give their orientations in terms of ϕ , δ , and λ ;
- find the angles Φ and Θ of the axes X , Y , Z , P , and T .

P20.3 For the 16 February 1998 Reykjanes earthquake ($M_w = 6.7$, lat. 52.68°N , long. 33.66°W), determine the fault plane solution.

- (a) Retrieve from the IRIS website the vertical component of seismograms for the following stations (see [Chapter 3](#)):

Station	Azimuth	Take-off angle
RES	334.30	33.70
KOG	205.10	29.10
TAM	119.70	31.20
EIL	86.60	27.70
SKD	89.60	30.90
SUW	64.50	33.50
FFC	301.10	31.90
ANMO	281.10	28.10
ANTO	78.40	30.20
CCM	273.20	31.30
COLA	330.50	28.30
KIEV	67.00	32.10
PTG	211.60	26.60
SDV	228.10	28.10
SJG	228.70	31.10
TIXI	6.60	27.30
MA2	357.50	23.30

- (b) Read the P wave polarity at each station.

- (c) Using the equal area projection (Schmidt net, EMP_20_1) or a computer program (websites given in the text) estimate the fault plane solution as in Problems [P20.1](#) and [P20.2](#).

P20.4 Using the spectrum of P waves of problem P3.5 (CART station for the earthquake of Mozambique, 22 February 2006) if $i_h = 24.3^\circ$, $i_o = 18.6^\circ$, the Poisson coefficient 0.25, the radiation pattern $R(\phi, \delta, \lambda, i) = -0.376$, $g(\Delta) = 0.41$, and $\alpha_h = 7.2 \text{ km s}^{-1}$, calculate the seismic moment, the radius of circular fault, the apparent average stress, and the stress drop.

In this final chapter, we apply the quantitative knowledge about earthquakes we have developed along the previous chapters to the subjects related to the occurrence of earthquakes, which we presented in [Chapter 2](#), as an introduction, in a descriptive form. First, we deal with the distribution of the occurrence of earthquakes in space and time and with their magnitude. This falls under the general term of what is called seismicity, at a global or regional level. Second, we present the relation of earthquakes and their mechanism with the dynamics of the Earth's lithosphere and plate motion or seismotectonics. Third, we address briefly the problem of assessment of seismic hazard and risk and of earthquake mitigation and prediction.

21.1 The spatial distribution of earthquakes

As we saw in [Chapter 2](#), earthquakes are distributed not on an aleatory form on the Earth surface but grouped at seismic active zones, while other zones are practically free of them. The development of modern, more reliable seismic instrumentation ([Chapter 3](#)) and methods for the determination of earthquake location, origin times, and size ([Chapter 16](#)) have led to a better understanding of seismicity, that is, the distribution of earthquakes in space, time, and magnitude. The global picture of earthquake locations shows that in great part they coincide with the plate boundaries ([Section 2.7](#), [Figs. 2.3](#) and [2.8](#)). Intermediate and deep earthquakes occur at subduction plate boundaries where the oceanic material of lithospheric plates is pushed into the mantle ([Section 2.7](#), [Figs. 2.4](#), [2.9](#), and [2.10](#)). As a general rule, at oceanic ridges there are only shallow earthquakes, whereas for island arc regions there are shallow, intermediate, and deep earthquakes ([Figs. 2.3](#) and [2.4](#)). Although shallow earthquakes are classified as those less than 30–60 km deep, most of them take place at depths less than 20 km, that is, in the upper rigid part of the crust that can break by brittle fracture which is called the *seismogenic* layer. Earthquakes at greater depths correspond to zones where this rigid and relatively cold material of the upper crust is introduced into the upper mantle by lithospheric subduction, where earthquakes may reach depths of up to 700 km ([Fig. 2.10](#)). Although most earthquakes are generated by brittle fracture of lithospheric material, very deep earthquakes may be produced by other processes related to phase changes in the focal material. According with their origin, earthquakes may be classified as *tectonic*, *volcanic*, or *anthropogenic* or *induced*. Most earthquakes that occur are due to tectonic processes, and rupture may be explained by crustal shear fractures. Volcanic earthquakes are originated through processes in the

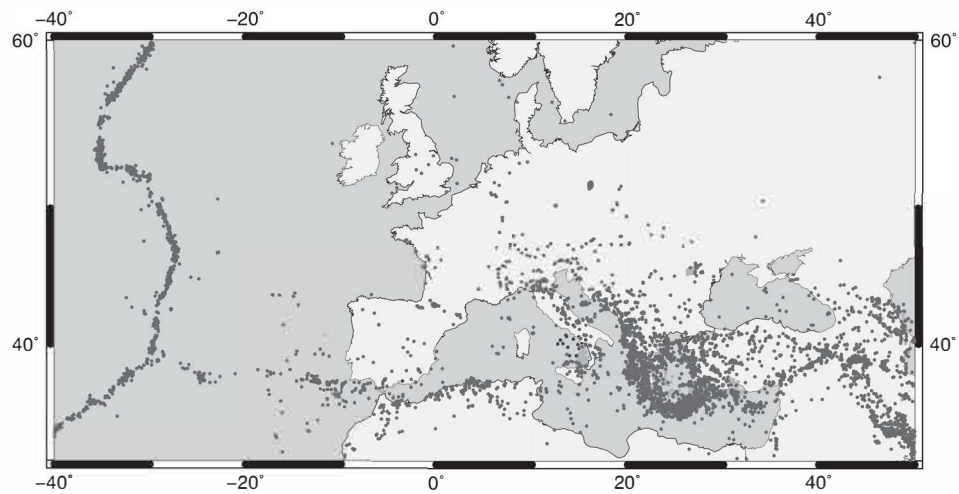


Fig. 21.1

Map of earthquakes in the Azores–Mediterranean region for the period 2000–2015 $M > 4$ (red shallow ($h < 60$ km), green intermediate depth ($60 < h < 300$ km), blue deep ($h > 300$ km)) (data from NEIC, U.S. Geological Survey). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

magmatic chamber, and their seismic source involves changes in volume. Anthropogenic earthquakes are caused by man's activity, such as mining, gas injection, fracking, dams, and reservoirs fillings, etc.

According to plate tectonics, the global distribution of epicenters is related to the relative motion between lithospheric plates (Figs. 2.3 and 2.8). Most earthquakes, then, take place at plate boundaries (*interplate*) and, less commonly, take place in plate interiors (*intraplate*). The most active regions in the world correspond to the margins of the Pacific Ocean. Earthquakes of large magnitude take place along this zone at its eastern part in the Americas, from the Aleutian Islands to southern Chile, and at the western part, from the Kamchatka peninsula in Asia to New Zealand. Besides shallow earthquakes, throughout most of this long region, at the boundary of the large Pacific plate, intermediate and deep shocks take place along the western margin of Central and South America and on the other side of the Pacific along the systems of island arcs (Aleutians, Kuril, Japan, the Philippines, New Guinea, Solomon, Fiji, New Zealand) (Fig. 2.4).

Another large seismically active region is known as the Mediterranean–Alpine–Himalayas region and extends from west to east from the Azores Islands to the eastern coast of Asia. This region is related to the boundary between the plates of Eurasia to the north and Africa, Arabia, and India–Australia to the south (Fig. 2.8). Its seismicity involves shallow, intermediate, and deep earthquakes (Figs. 2.3 and 2.4). A third seismic region is formed by earthquakes located at ocean ridges that form the boundaries of oceanic plates, such as the Mid-Atlantic Ridge, East Pacific Rise, etc. (Fig. 2.3). In these regions earthquakes are of shallow depth and are concentrated in relatively narrow bands following the trend of the oceanic ridges (Fig. 2.3). On comparing Figs. 2.3 and 2.4, we can see the location of shallow and intermediate and deep earthquakes and the presence of zones of

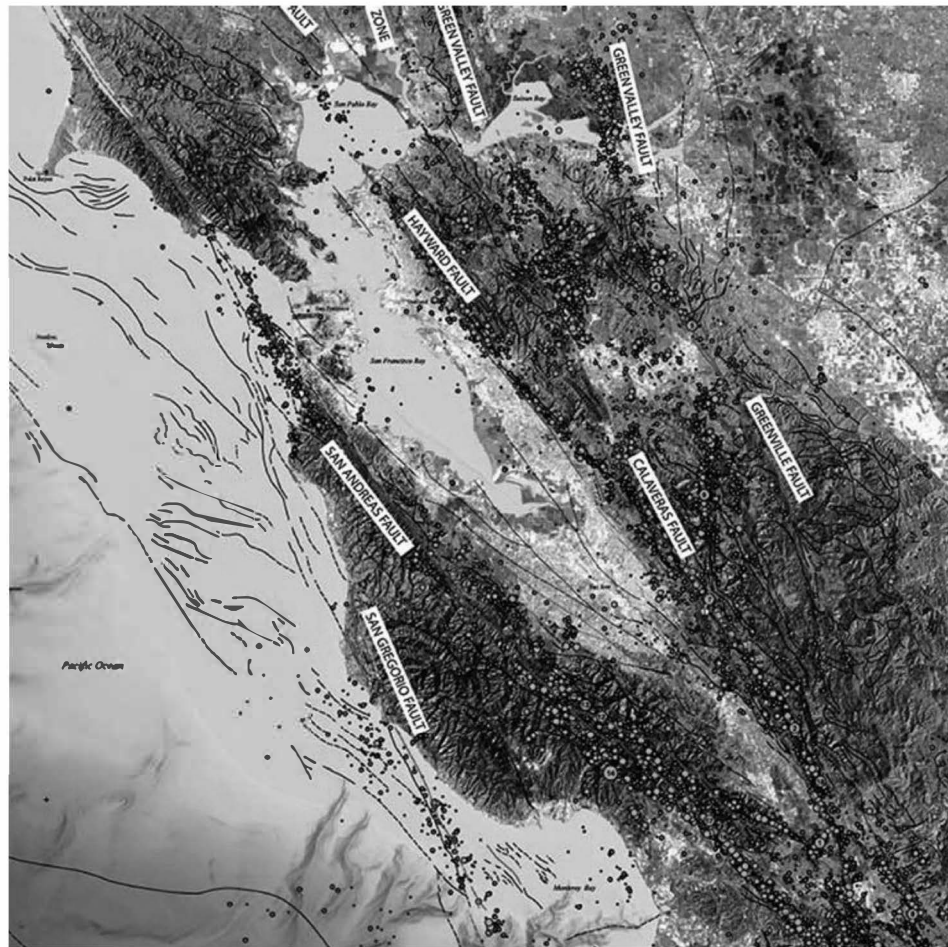


Fig. 21.2

Geological faults and earthquake distribution in California (with permission of U.S. Geological Survey). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

concentrated seismicity, and others with earthquakes spread over wide areas. In general, boundaries between oceanic plates and between oceanic and continental plates have simpler distributions of seismicity than do boundaries between continental plates.

An example of a complex distribution of seismicity at a boundary between two continental plates is the Mediterranean region (Fig. 21.1) (Udías and Buforn, 1994). Inside this region the most active zones in order of activity correspond to Greece, Turkey, Italy and the Alps, the Carpathians and Dynarics, North Africa, southern Spain and the Pyrenees, and the Rhine Graben. Along this region, seismic activity is formed by a continuous occurrence of earthquakes of small and moderate magnitude ($M < 6$) with the occurrence of large earthquakes ($M > 6$) separated by larger intervals of time. Destructive earthquakes have happened more or less often in all active zones of these regions. Most earthquakes are of shallow depth, with four zones of intermediate and deep earthquakes, namely, the Hellenic arc, the Carpathians, Sicily–Calabria, and southern Spain and northern Morocco.

The deepest earthquakes have taken place in southern Spain (640 km) and Sicily–Calabria (450 km). In the other two regions (Hellenic arc and Carpathians) the maximum depth is about 200 km. As can be seen from Fig. 21.1, earthquakes are spread over wide areas, indicating the presence of several small plates between the two large plates of Eurasia and Africa and of zones of diffused seismicity.

At a regional scale, earthquakes are distributed more or less related to the presence of active geological faults. A well-known example is that of earthquakes in California (Fig. 21.2). Earthquakes that break the surface can be easily related to faults but this is not often the case. To assign earthquakes to specific geological faults it is necessary to have a very accurate epicentral determination.

21.2 The temporal distribution of earthquakes

As we already mentioned in Chapter 2 (Section 2.4), from a statistical point of view, the simplest model of the occurrence of earthquakes with time is a *Poisson distribution*, which assumes that earthquakes are independent events. If λ is the rate of occurrence of earthquakes within a time t , the probability that n earthquakes take place within such an interval is

$$p(n) = \frac{\lambda^n e^{-\lambda}}{n!}. \quad (21.1)$$

If the occurrence of earthquakes follows a Poisson distribution, then the intervals of time δt between consecutive earthquakes have an exponential distribution. The probability that two earthquakes are separated by an interval δt is

$$p(\delta t) = \lambda e^{-\lambda \delta t}. \quad (21.2)$$

For a given rate of occurrence, the probability is greater for smaller intervals. As already mentioned in Chapter 2, studies of temporal series of earthquakes show that they deviate, to a greater or lesser extent, from Poisson's law. The interdependence of earthquakes shows up in the occurrence of clusters such as sequences of *aftershocks* and *swarms*.

As mentioned in Chapter 2, the most important clustering of earthquakes is that of the occurrence of aftershock series after large earthquakes. The distribution in time of the number of aftershocks follows, generally, an inverse power law of time. In 1894, Omori proposed a dependence of $1/t$, which has been modified to the form (Utsu, 1961)

$$N(t) = \frac{K}{(c + t)^p}, \quad (21.3)$$

where K and c vary with each series and p has values in the range 0.7–1.4. Also here, values of p are interpreted in terms of the heterogeneity of the fault material. The number of aftershocks diminishes more rapidly for a very homogeneous material (high values of p) than it does for a heterogeneous material (low values of p).

An interesting phenomenon in the occurrence of earthquakes is that of so-called *earthquake doublets*, when an earthquake is followed after a short time interval by another shock of similar magnitude located very near to it. The term earthquake doublet is not clearly defined and different mechanisms have been proposed for its occurrence, for example, triggering of the second event by increases in the *Coulomb stress transfer* produced by the occurrence of the first.

21.3 Seismic cycles

Earthquakes occur when the stress overcomes the strength of the source material at a critical value, producing its fracture, or in the friction model when the applied stress overcomes the static friction. The accumulation of stress in a particular area is produced by tectonic processes that are related ultimately to plate motion. If we assume that stress accumulates at a constant rate, it will reach the critical value and produce earthquakes in a certain fault area at more or less regular intervals. The process of the accumulation of stress and its release by the occurrence of an earthquake on a particular fault constitutes a *seismic cycle*.

In the simplest case, the rate of accumulation of stress is assumed to be uniform and reaches the same critical level at which an earthquake is produced with constant drop in stress, reducing the stress to the same level, from which it starts to accumulate again. In this ideal case, all earthquakes would have the same magnitude and repeat at constant intervals with periodic seismic cycles. Then, the magnitude and time of occurrence of earthquakes could be predicted (Fig. 21.3a). A more realistic model was proposed by Kunihiko Shimazaki and Takashi Nakata (1980), whereby stress is released when it reaches the same maximum value, but the drop in stress for earthquakes is not always the same; that is, stress is released by earthquakes of different sizes. The intervals between earthquakes are not always the same, but rather depend on the size of the preceding earthquake (Fig. 21.3b). This situation, known as the *time-predictable model*, allows prediction of the time of occurrence of the next earthquake, but not its size. A model with earthquakes happening when the stress reaches various levels and also with different drops in stress will have earthquakes of different sizes occurring at different intervals. In this case, we cannot predict

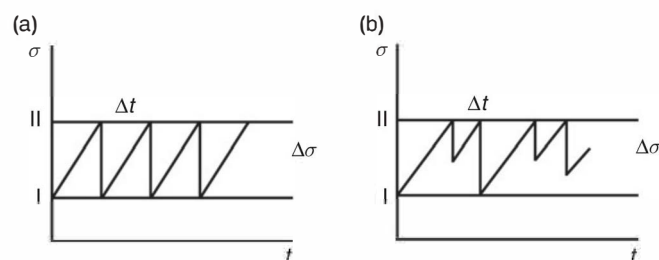


Fig. 21.3

Models of the occurrence of earthquakes: (a) periodic, equal drop in stress, and equal time interval; (b) non-periodic, different drops in stress, and different time intervals (the time-predictable model) (Shimazaki and Nakata, 1980) (with permission from the American Geophysical Union).

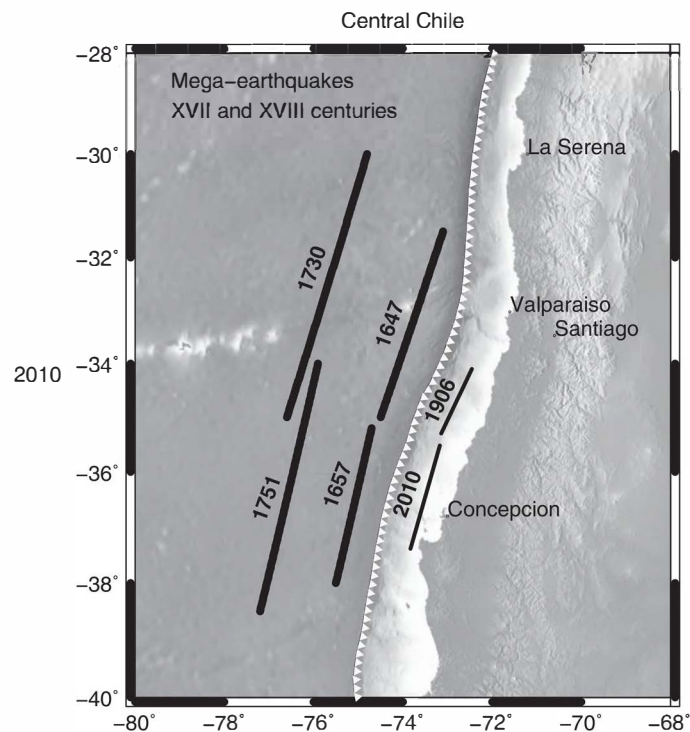


Fig. 21.4 Length of faults in central Chile broken by the large earthquakes of 1647, 1657, 1730, 1751, 1906, and 2010 (courtesy of R. Madariaga).

the time or magnitude of the next earthquake from knowledge of the preceding seismic cycle.

The concept of a seismic cycle is very important if one is to understand the occurrence of earthquakes on the same fault. Although the accumulation of stress may be approximately uniform, the heterogeneity of the material does not allow the existence of periodic or regular cycles of occurrence of earthquakes of the same magnitude. Owing to the heterogeneous conditions at the fault, stress may be released starting at various levels and producing earthquakes of different sizes, so that the time-predictable model is not valid. For example, in central Chile two large earthquakes broke the same fault in 1647, and 1657, and again one in 1751, to break again in 1906 and 2010 (Fig. 21.4). The time intervals between earthquakes (10, 73, 155, and 104 years) and also the sizes of the earthquakes were different. Another problem that affects the regularity of seismic cycles is the existence on a fault of *aseismic slip* caused by creep phenomena. Thus, part of the accumulated stress is released by continuous aseismic slippage of the fault, independently from the occurrence of earthquakes. The aseismic slip depends on the nature of each fault and it is difficult to observe. The result is that the sum of the stresses released by earthquakes does not add up to that accumulated by tectonic processes and so the seismic slip is only part of the total slip on the fault.

A concept related to the seismic cycle is that of the *characteristic earthquake* (Schwartz and Coppersmith, 1984; Aki, 1984). This concept is based on the consideration that a particular fracture, owing to its dimension remaining constant, can only produce earthquakes with a certain maximum size. For each fault, earthquakes with this maximum magnitude are called characteristic earthquakes. These earthquakes have very similar characteristics and repeat themselves on the same fault, at quasi-regular intervals. Implicit in this concept is that each fault is an independent unit that breaks in its totality, producing always an earthquake of the same maximum magnitude. However, it has been observed that large earthquakes may break through several faults, and most often earthquakes do not break the whole length of the fracture, thus, the principle of characteristic earthquakes is not always satisfied. Many do not even accept the validity of the concept itself. Here again, the problem is that the conditions in fault regions are very different from each other and are not homogeneous, thus regularities in the occurrence of earthquakes valid for all cases are difficult to establish.

The main problem with respect to the regularity of seismic cycles and characteristic earthquakes is that observations of the occurrence of large earthquakes are very limited in availability. Instrumental data started to be recorded at the beginning of the twentieth century and reliable global data do not extend back more than some 50 years. Historical earthquakes dating back several centuries may be known in some regions, but our knowledge of them has many limitations. Even admitting the reliability of data of this type, the record cannot be extended back any further than about 2000 years ago. In most active regions, large earthquakes ($M > 6.5$) repeat at intervals that vary from tens to hundreds of years (as we have seen for central Chile, Fig. 21.4). In regions of low seismicity these intervals may be larger, of the order of hundreds of years. Moreover, very large earthquakes ($M > 8$) are very rare events that may be separated by even greater intervals. For example, it is difficult to know the return time for the great Lisbon earthquake of 1755 or the San Francisco earthquake of 1906. The recent Haiti 2010 and Tohoku 2011 earthquakes with large numbers of casualties came as a surprise to seismologists.

A new source of observations about earthquake occurrence comes from *paleoseismicity* studies of active faults. Detailed field observations of the slip on trenches through faults allow the dating of large earthquakes in the geological past, but there are still only a few of these studies. Moreover, dating these events is difficult and individual faults may cease at some time to be active. Modern systematic observations of coseismic and aseismic deformations on active faults by InSAR and GPS satellite geodetic methods give further information on earthquake cycles, but these are still recent sources of information. In conclusion, with the information we have at present, it is still difficult to solve the problem of the periodicity of large earthquakes and to know to within a certain precision the characteristics of seismic cycles.

21.4 The distribution of magnitudes

Observations show that, in seismic regions, during any period of time, the number of small earthquakes is many times that of larger ones. This fact is expressed in the empirical law

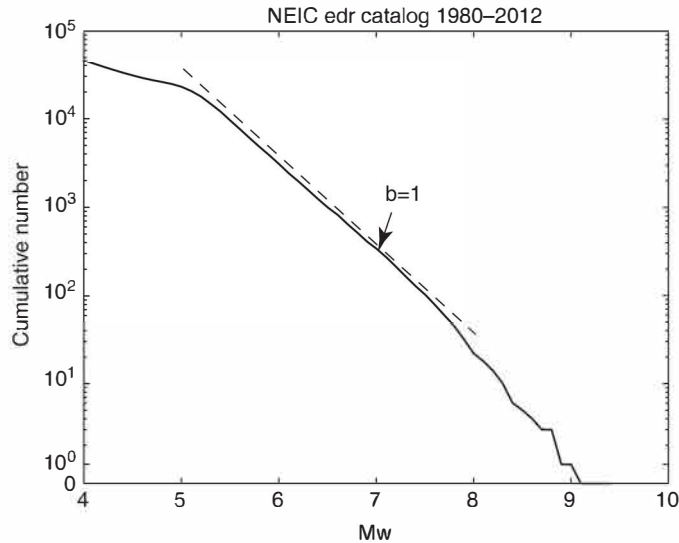


Fig. 21.5 The cumulative number of earthquakes with magnitude.

suggested by Omori in 1889 and proposed by Gutenberg and Richter (1956) in logarithmic form (*Gutenberg and Richter law*):

$$\log N(M) = a - bM, \quad (21.4)$$

where $N(M)$ is the number of earthquakes of magnitude larger than M , a is a constant that represents the number of earthquakes of magnitude larger than zero, and b is the proportion of earthquakes with small and large magnitudes. Although we do not know the exact form in which the elastic energy accumulated in a region is released by earthquakes, the distribution of their magnitudes follows this universal law. The value of b is an important parameter in studies of seismicity. It is determined by finding the slope of the straight-line fit to the distribution of the logarithm of the number of earthquakes versus their magnitudes. In general, observations fit this distribution, with some deviations, corresponding to the lack of knowledge of the very small and very large earthquakes (Fig. 21.5). For a large area and long interval of time, these deviations are generally due to a lack of completeness in observations at both ends of the magnitude range. For observations from a single fault zone, deviations may be related to its conditions and the size of the largest earthquakes.

Values obtained for b for very large samples of data are very stable, in the range 0.6–1.4, and its most common value is very near to unity. If the samples are small, say for a region of low seismicity, values of b , strongly deviating from unity, are not significant. Changes in values of b on going from one region to another are sometimes related to their mechanical characteristics. High values of b indicate a high number of small earthquakes, which is to be expected in regions of low strength and large heterogeneity, whereas low values indicate the opposite, namely high resistance and homogeneity. Changes of b with time for the same region have been associated with changes in stress conditions, and, hence, b has been proposed as a parameter to consider in the problem of the prediction of earthquakes. In both

cases, values of b have to come from a very large number of earthquakes if their changes are to be significant. For this and other reasons, the degree of significance of the change of b has been questioned by some authors.

21.5 Models of the occurrence of earthquakes

Observations of the occurrence of earthquakes in space, time, and magnitude have led to the proposal of models that simulate certain properties of their distribution. These models specify the form in which accumulated elastic stress is released along the fault, producing earthquakes of varying magnitude at different times. Models describe the process of stress loading in terms of elastic strain in the focal region and its subsequent release with the occurrence of earthquakes. As we have seen, the process of stress loading and the subsequent release of stress constitutes a seismic cycle. Models must satisfy Gutenberg and Richter's relation and that the universal value of b is near unity, and should try to explain this fact.

The first model proposed to simulate the properties of the occurrence of earthquakes consists of a series of blocks, connected to one another by springs that slide on a horizontal surface with friction, as they are drawn by being connected to a large block that moves with a constant velocity (Fig. 21.6) (Burridge and Knopoff, 1967). The moving block represents the material on one side of a fault that is drawn by tectonic movements loading springs that connect the small blocks. Owing to friction between blocks and the stable surface, motion is not transmitted directly to the blocks, but rather first deforms the springs. Once the springs have accumulated enough energy to overcome the friction of the blocks, these move, sliding on the stable surface. When a block moves, it loads elastic energy onto the springs that link it to adjacent blocks, causing their later displacement. The motion of one or a few adjacent blocks represents small earthquakes and that of many blocks corresponds to a large one. The smallest earthquake is represented by the motion of one block and the maximum size possible is represented by that of all blocks of the system. Owing to friction, the blocks do not move uniformly, but rather by sudden increments, in the form called *stick-slip motion*.

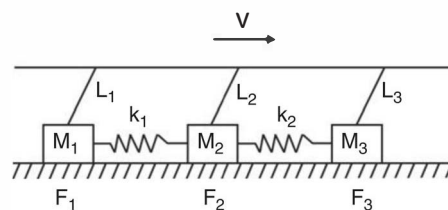


Fig. 21.6 The model of blocks and springs for simulating the occurrence of earthquakes.

With models of this type, we can simulate the occurrence of earthquakes with a proportion of small and large earthquakes similar to that of Gutenberg and Richter's law. Another result of these models is that small events take place in a random manner whereas large ones occur in a quasi-periodic way. Originally, the model was only one-dimensional, but it has been generalized to two and three dimensions. More recent models are more complex with viscoelastic elements between blocks, the friction variable with the velocity, and other conditions that allow the simulation of series of aftershocks after large events and other properties of the occurrence of earthquakes (Dieterich, 1972; Mikumo and Miyatake, 1979). The fundamental idea of these models is that each element of a fracture interacts with those surrounding it and thus influences the evolution of the whole system. Although these systems are in themselves deterministic, they behave chaotically, as their complexity is increased, due to the increase in their instability. The result is a quasi-chaotic behavior in which very small changes in initial conditions lead to very large effects on the evolution of the system which behaves more and more randomly with time.

Another approach to studying the occurrence of earthquakes is to consider their *fractal* nature, using the fractal theory developed by Benoît Mandelbrot (1977) (Andrews, 1980; Aki, 1981; Hirata, 1989). A basic property of all fractal distributions of a variable r is that they obey a power law of the type Ar^{-D} , where the exponent D represents the *fractal dimension* and is related to the coordinates associated with the elements of the set that represent the phenomenon being studied. The Gutenberg and Richter law (21.4) can be transformed into a *power law* for the number of earthquakes as a function of the seismic moment $N(M_0)$ by substituting equation (16.44) into it:

$$N(M_0) = BM_0^{-D}, \quad (21.5)$$

where B is the constant for a constant drop in stress. Since the seismic moment is proportional to the source dimensions, a similar relation can be written for the fault's length or area using (16.39) and (16.42). According to Aki (1981), $D = 2b$ and, since b has a value near to unity, D , the fractal dimension, is approximately equal to two. This result can be interpreted as representing the fact that earthquakes are distributed over two dimensions (a plane), which agrees with the observation of their occurrence on fault planes.

The fractal nature of the occurrence of earthquakes implies that they have a *self-similar* stochastic distribution, that is, they behave in a similar form, independently from the range of sizes considered. Although this is the general rule, *self-similarity* need not be perfectly satisfied. It has been observed that the relation between fault length and the seismic moment is different for moderate and small earthquakes ($M < 6.5$) than for large ones ($M > 6.5$) (Scholz, 1982). In the first case the seismic moment is proportional to the cube of the source length, whereas in the second it is proportional to its square. This break in the self-similar rule can be explained by invoking the influence of the thickness of the seismogenic layer. As we saw, this layer has an approximate thickness of 20 km and thus affects the geometry of the seismic source, which is nearly circular ($L = W$) for small earthquakes and rectangular ($L > W$) for large ones (Section 17.1).

Once we have accepted the fractal nature of earthquakes and their self-similarity, we may ask what dynamic process is responsible for this behavior. Some authors have answered this

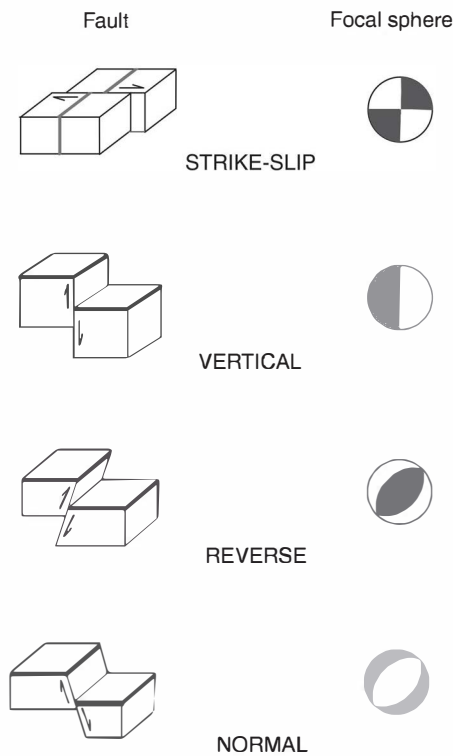
question in terms of *self-organized criticality* (SOC) (Bak and Tang, 1989; Sornette and Sornette, 1989; Ito and Matsuzaki, 1990). This term (SOC) is applied to dynamic systems that evolve by themselves until they reach a critical state in which phenomena take place randomly, according to a power law. In the case of earthquakes, the system is the material of the Earth's crust, which evolves under tectonic stresses until it reaches a state of SOC. In this state, earthquakes of all sizes take place with only the limitation of a minimum possible size (a minimum earthquake) and a maximum size imposed by the dimensions of the seismogenic structure (the so-called maximum or characteristic earthquake). This behavior is a consequence of the non-linear dynamic space-time characteristics of the Earth's crustal system in response to stresses generated by lithospheric plate motion.

This type of process can be modeled, in a form similar to that of the models with blocks and springs, by a distribution of *cellular automata* (developed by John von Neumann to solve very complex problems in the 1940s) whose behavior under an applied stress is specified (Barriere and Turcotte, 1991; Georgudas *et al.*, 2007). In this type of model, each cell receives and transmits stress to its neighbor cells as it moves, changing the stress distribution of the whole system. After a sufficiently long time, the system evolves by itself into a critical state (SOC), in which events of all sizes are produced with a power-law distribution. The size of events is given by the number of adjacent cells that move at a given moment. Models of this type reproduce certain properties of the occurrence of earthquakes, such as the value of b being near unity and the occurrence of aftershocks that decay in number with the inverse of time after large earthquakes.

21.6 Seismotectonics

As we saw in Section 2.7, seismologic data used in seismotectonics include the geographic distribution of epicenters, with indications of their magnitude, depth, and their focal mechanism. A very common representation of the focal mechanism in seismotectonic studies is shown by means of the projection of the focal sphere with quadrants of compressions and dilations with different color. The correspondence of diagrams of this type to various types of faults (strike-slip, vertical, normal, and reverse) is shown in Fig. 21.7, together with the projection of the focal sphere onto the horizontal plane. In strike-slip faults the two planes that separate the quadrants are vertical, whereas in normal and reverse faults, the center of the projection is in a dilation and a compression quadrant, respectively. Another graphical representation of focal mechanisms is by horizontal projection of slip vectors and of pressure and tension axes. Slip vectors are related to kinematic aspects of tectonics, that is, the directions of motion of plates or blocks. Pressure and tension axes refer to dynamic aspects, that is, stress orientations in a particular region.

As we saw in Section 2.7, plate boundaries can be reduced basically to three types, namely, divergent or rift zones, convergent or subduction zones, and transcurrent horizontal slip or transform faults (Fig. 2.9). At divergent boundaries, earthquakes are of shallow depth and moderate magnitude ($M < 7$), forming a narrow band along boundaries (for

**Fig. 21.7**

The correspondence between faults and diagrams of the focal mechanism (horizontal projection of the focal sphere). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

example on the Middle Atlantic Ridge). Their focal mechanisms correspond to normal or strike-slip faulting with the horizontal tension axis normal to the plate boundary. At convergent boundaries earthquakes reach large magnitudes ($M > 7$), and their mechanisms for shallow shocks correspond to thrust or reverse faulting due to horizontal pressure normal to the trend of the boundary (for example Pacific Ocean boundaries). Deep earthquakes have mechanisms with either the pressure or the tension axis in the direction of the subducted plate, indicating its condition under compression or tension. Generally, in the upper part, tension axes are along the subducted plate while in the lower part are the pressure axes. Thus, the upper part of the plate is stretched while the lower part is compressed against the mantle. Along the transformed faults that connect two extension, two subduction or an extension and a subduction boundary, earthquakes are shallow with a strike-slip mechanism and may reach large magnitude ($M > 7$) (for example, San Andreas fault). The mechanisms of shallow earthquakes on the convergence front are of reverse or thrust faulting corresponding to the horizontal pressure being normal to the front and of strike-slip motion corresponding to pressure along the transform fault (Fig. 21.8a). For two zones of extension connected by a transform fault, the mechanism of earthquakes in the rift zone corresponds to normal faulting, whereas that in the transform fault corresponds to strike-slip faulting (Fig. 21.8b).

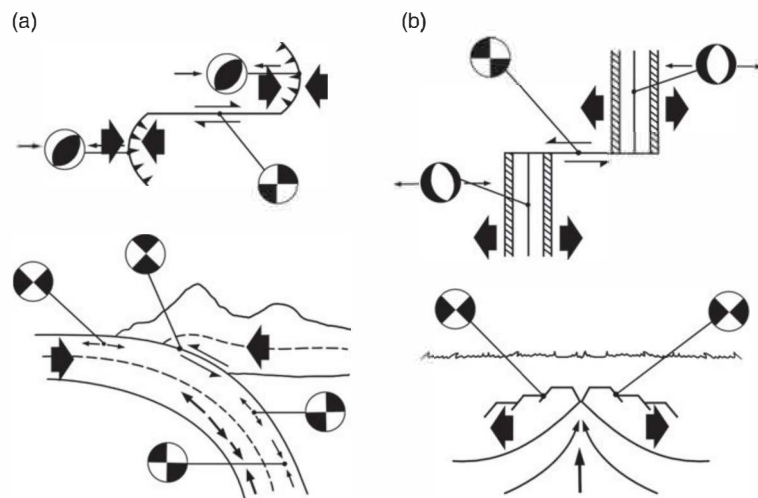


Fig. 21.8

The connection of convergent (subduction) and divergent (ridge) zones by means of transform faults and types of earthquakes focal mechanisms: (a) to subduction zones. (b) Two ridges.

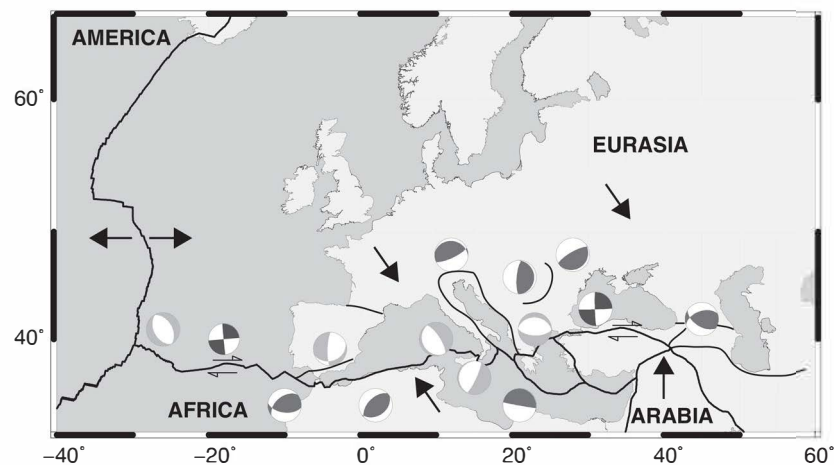


Fig. 21.9

The seismotectonic framework for the Azores–Mediterranean region. (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

Figure 21.9 shows an example of a seismotectonic framework for the Mediterranean region based on seismicity and focal mechanism data (Udías and Buform, 1994). The tectonic situation is due to the converging motion of the plates of Eurasia to the north and Africa and Arabia to the south. The situation is complex due to the continental nature of both plates and the presence of several small plates, such as the Italy–Adriatic, Aegean, and Anatolian plates. The seismicity is somewhat diffused over a wide region, indicating that there are some intraplate deformations. The predominant types of focal mechanisms are shown together with the direction of horizontal stresses and slip directions.

The regional stresses correspond to horizontal pressure in the NW–SE to N–S direction due to the converging motion of Eurasia, Africa, and Arabia, with two localized regions of horizontal tensions, in the NE–SW direction in central Italy and in the N–S direction in northern Greece. Two regions of right-lateral strike–slip motion are located in the Azores–Gibraltar and north Anatolia faults. Two subduction zones are located in the Hellenic arc and Sicily–Calabria, with another two zones of intermediate and deep earthquakes in southern Spain–northern Morocco and the Carpathians.

21.7 Seismic hazard and risk

As was mentioned in [Section 2.8](#), we differentiate between seismic risk and hazard. The seismic hazard of a site is given in terms of the ground motion (displacement, velocity, and acceleration) produced by nearby earthquakes independent of the damage produced. Risk involves the damage produced incorporating the vulnerability of the buildings. In seismic hazard studies, commonly, only accelerations are considered, and, in many instances, ground motion is represented by values of the seismic intensity. The ground acceleration is the parameter that is most directly related to engineering design, so modern seismic hazard maps are given in terms of this parameter. However, often ground accelerations are derived from seismic intensities and this relation has a wide margin of variability.

The seismic hazard can be defined and evaluated in a *deterministic* or a *probabilistic* form. The first is based directly on the previous occurrence of earthquakes that have affected a given site and the second as the probability that ground motion of a certain amplitude takes place at a given site during a certain period of time using past earthquakes and certain probability distributions. Notice that seismic hazard is not seismic prediction. The seismic hazard of a site depends on many factors, namely, the distribution of nearby seismic sources, their magnitudes and focal mechanisms, the nature of the propagating medium, especially wave attenuation and geometric spreading, and characteristics of the shallow layers under the site that increase or decrease amplitudes of ground motion.

Deterministic methods are based on the assumption that the seismicity near a site in the future will be the same as that in the past and they are known as *deterministic seismic hazard assessment* (DSHA) (Panza *et al.*, 1999). Thus, the seismic hazard is given by the maximum values of amplitudes of ground motion that have been produced by the largest earthquakes in the past. The largest potential earthquake affecting an area is often called the *maximum credible earthquake* (MCE). Then, the better we know the seismic history of a region, the better we can determine its hazard. These methods depend very strongly on our knowledge of the past seismicity of the region surrounding a given site. The greatest problem is that often we do not know with great accuracy about the occurrence of earthquakes over a very long period of time. This is more crucial for regions of moderate seismicity where large earthquakes are separated by hundreds of years. It makes sense, however, to be prepared for repetition of the strongest earthquake that we know has

happened in the past, as is recommended by the International Seismic Safety Organization (ISSO) (Mualchin, 2010).

Probabilistic methods use statistical distributions to represent the seismic activity of a region on the basis of the occurrence of earthquakes in the past and are widely used and especially endorsed by the U.S. Geological Survey; they are known as *probabilistic seismic hazard assessment* (PSHA) methods (Cornell, 1968; Merz and Cornell, 1973). The result is determination of the probability of the occurrence of a certain level of ground motion or seismic intensity for a particular time interval. The occurrence of earthquakes in the future is not given directly by that in the past, but rather is estimated from a statistical distribution. In order to calculate the expected ground motion, earthquakes of the past are not considered in terms of their exact location, but rather are assigned to established *seismogenic zones* that group earthquakes in areas supposed to have the same tectonic origin. The determination of these zones (*seismic zonation*) is an important problem in the determination of seismic hazards. Different zonations may lead to different estimations of hazards from the same set of earthquake data. Probabilistic methods provide estimations of probability for various levels of intensity, rather than solely maximum values as do some deterministic methods. The seismic hazard in probabilistic terms is, then, given by the probability that, in a particular region and within a given time interval, ground motion reaches a certain amplitude or the seismic intensity attains a certain degree. This can be formulated also as the probability that, during a certain time interval, the maximum value of the intensity or ground acceleration exceeds a particular value. The inverse of this probability, or the time in which we expect with a certain probability that a certain intensity (or ground acceleration) will be exceeded, is called the *return period* for that intensity or ground acceleration. This is determined from knowledge of earthquakes in the past that have happened at various distances from its neighborhood according to the established zonation.

Since seismic intensities are used in most hazard studies, we need to know the laws governing the attenuation of intensity with distance (Section 16.3, Fig. 16.3). Generally an expression similar to (16.10) is used, whereby the values of the coefficients a and b are determined from intensity maps for earthquakes in the region. With the availability of dense networks of accelerographs, we can determine directly the attenuation of the ground acceleration with distance. From knowledge of the occurrence of earthquakes over time and their maximum intensities, inside each seismogenic zone surrounding a site, together with the laws governing the attenuation with distance, we can determine their expected ground motion or intensities at the site where we calculate the seismic hazard (Fig. 21.10). In probabilistic methods we can calculate the probability of a certain level of intensity (or ground acceleration) being exceeded within a given time interval that defines the seismic hazard; for example, the probability that intensity VIII (or a ground acceleration of 0.4 g) is exceeded within 200 years. If we calculate values of the seismic hazard at many sites we can draw maps (known as PSHA maps) with intensities or ground accelerations expected with a given probability within a certain interval of time (Fig. 21.11). Recent earthquakes with ground accelerations well above those predicted by PSHA maps, for example, those in Italy in 2009 (L'Aquila) and 2016 (Amatrice), and the disastrous earthquake of Port-au-Prince (Haiti) in 2010 with 316,000 casualties, have questioned the results of these methods, creating a debate between the convenience of using deterministic and probabilistic methods

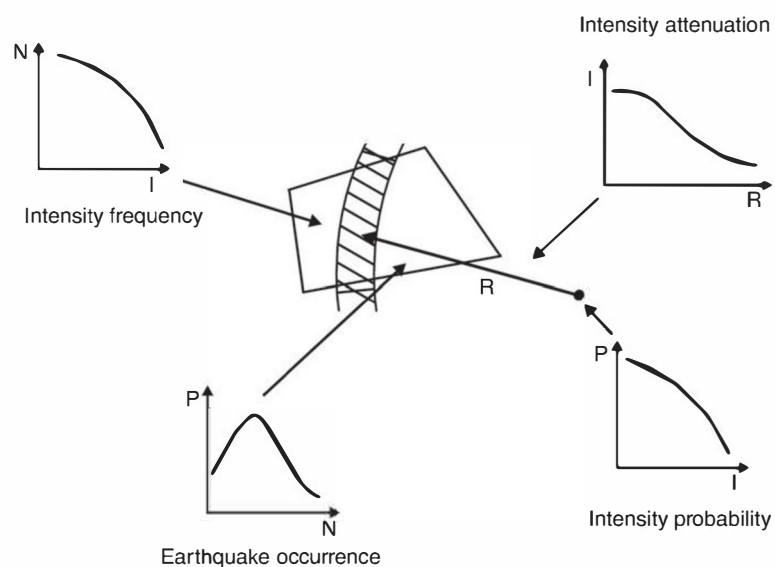


Fig. 21.10

A schematic representation of the determination of seismic hazard (courtesy of D. Muñoz).

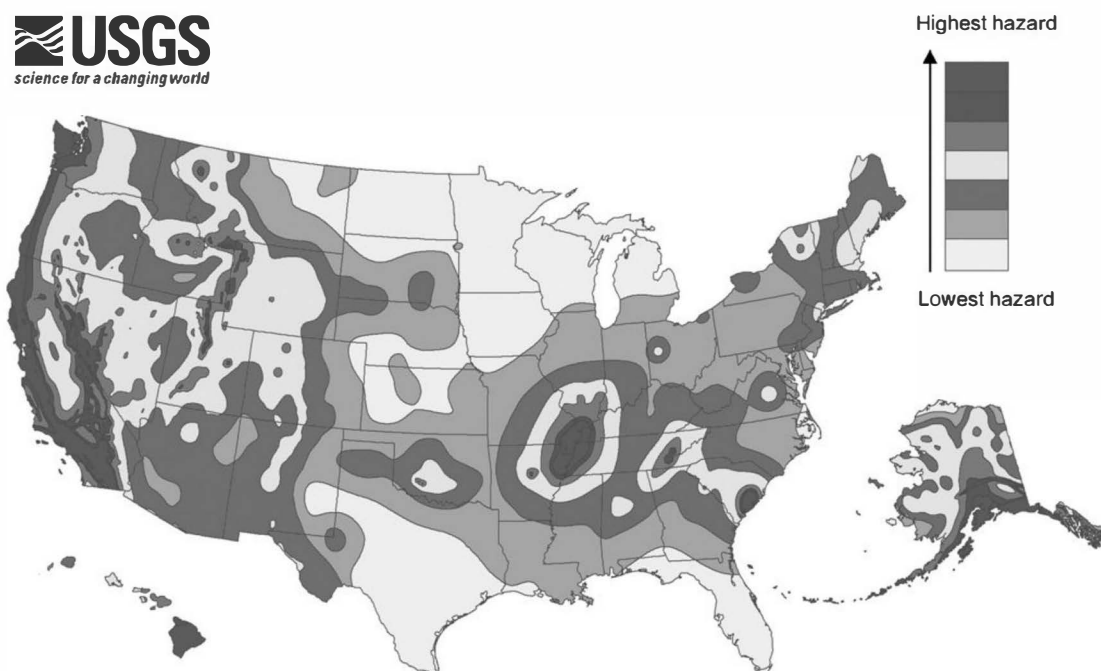


Fig. 21.11

A seismic hazard map for the USA (with permission of U.S. Geological Survey). (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

(Bommer, 2002; Zuccolo *et al.*, 2011). Ultimately all methods have their advantages and disadvantages, they must be used with consideration of their limits and they ultimately depend on the accuracy of our knowledge of the seismic history for a given region.

Adequate assessment of seismic hazards is fundamental to prevention of the damage caused by earthquakes, since buildings and other structures can be designed to withstand, without collapsing, the maximum ground acceleration expected for a given site. As we saw in the intensity scale (Table 16.1), for a given intensity, the damage to buildings depends greatly on the building materials and structural design which define their vulnerability. For intensities of about VIII, many masonry buildings of adobe or rubble stones suffer very heavy damage and even total destruction, but only a few buildings of reinforced concrete with a minimum of antiseismic design suffer moderate damage. Reinforced concrete and steel structure buildings with high levels of anti-seismic design can escape very heavy structural damage even at maximum levels of intensity. The study of the structural responses of various types of buildings is fundamental to safe building practice in seismic regions. All countries with an appreciable seismic risk have anti-seismic building codes that are sanctioned by law. Safety norms are stricter for structures whose collapse can produce large catastrophes, such as nuclear plants, dams, hospitals, public buildings, etc. If these norms are maintained in building, damage can be drastically reduced and total collapse avoided, with a consequent reduction in numbers of casualties. The problem remains for old buildings and inadequate construction in developing countries.

In recent years, a new tool has been developed to mitigate the damage of earthquakes: *Earthquake Early Warning Systems* (EEWS). An EEWS is a real-time system able to detect an earthquake after the first waves are recorded at a nearest station, so that fast notification of its potential to cause damage in a target area can be given before the destructive waves arrive (Wu and Kanamori, 2005; Allen *et al.*, 2009). The hypothesis of an EEWS is that the initial portion of the P wave contains information on the size of the earthquake, allowing determination of the final size before the arrival of S waves, which cause the greatest damage. Warning times range from a few seconds to a little more than a minute and are primarily a function of the distance of the user from the earthquake epicenter. At the present time, EEWS are operating in some countries, such as Japan, Mexico, and Taiwan, and are under study in others, such as Italy, Turkey, and Spain.

21.8 The prediction of earthquakes

As mentioned in Section 2.8, prediction of earthquake occurrence remains always an unrenounceable aim for seismologists (Hough, 2010). A true prediction means specifying the location, time, and magnitude of an impending earthquake to within narrow margins and with a high level of probability, with sufficient advance warning so that certain measures to reduce its effects can be taken. Prediction can be considered in long, medium, and short terms. Long-term prediction, in practice, can be identified with the assessment of the seismic hazard of a region. In this context, the determination of a high probability for the

occurrence of a large earthquake in an area does not constitute a true prediction. Medium- and short-term predictions are really related. For example, a medium-term prediction could be that an earthquake of magnitude about 6 will happen in a certain area within 4–6 months. As this time approaches, prediction requires that the probability increases and the time and place of occurrence are specified more precisely. Finally, enough evidence should make it possible to give a short-term prediction, say a few days before, specifying the size, place, and time of the impending earthquake within relatively narrow margins, so that the authorities can take adequate measures for protection of the population. In the actual state of the problem, there is not yet a methodology that allows medium- and short-term predictions of this type.

The problem of prediction is based on the identification of *precursor phenomena* that indicate the impending occurrence of an earthquake. We have already mentioned (Section 2.8) that strictly speaking, although it is related since antiquity and is present in popular opinion, the abnormal behavior of animals cannot be considered a true seismological precursor. Although it is often said that animals have sensed the earthquake before it happens, it is difficult to verify and quantify this assertion, let alone establish their relation to the earthquake. Generally, it is something established after the earthquake has happened. Seismological precursors are observables that are related to changes in physical conditions in the focal region before the occurrence of the earthquake. Precursors can be divided into seismic and non-seismic.

One seismic precursor is the observation of patterns of seismicity in a region with variations in space and time. This observation has shown the existence of *seismic gaps*, areas in active regions where there is low seismicity during a certain time previous to a large earthquake. In an active region, where the seismicity is expected to be distributed more or less homogeneously, the existence of an area where the level of seismicity is lower than average is called a seismic gap and is considered the most probable place for a large earthquake in the future. The size of the seismic gap and the duration of the lack of activity can be correlated to the magnitude of the future large shock. Correlations between the existence and extent of seismic gaps and the occurrence of large earthquakes have been established *a posteriori* in several instances. However, the use of this type of precursor as a predictor is not very reliable, because it is not known when a large earthquake will be produced in an observed seismic gap. A second seismic precursor is the observation of changes in a given active region of the rate of occurrence of small earthquakes, or background seismicity, with intervals of decreasing or increasing activity. In the first case (decreasing activity), we have a *seismic quiescence* that is associated with an impending larger earthquake. Also the opposite effect, an increase in activity, can be associated with a future earthquake. Although it may seem contradictory, both effects can be considered precursors; that is, a change in the rate of occurrence of small shocks may signal the occurrence of a large earthquake. For example, the rate may first diminish and later increase, or vice versa. Just like in the previous case, this type of change in seismicity preceding an earthquake has been observed *a posteriori*, but its use as a true precursor is problematic. For example, the L'Aquila (Italy) earthquake of 2009 ($M_w = 6.3$) was preceded by the increasing occurrence of smaller earthquakes over three months, but it could not be said with certainty that a larger one was to occur as happened.

Another seismic precursor that had wide acceptance at one time is the change in the velocity of seismic waves in the neighborhood of the focal region, previous to the occurrence of an earthquake. In some cases, it has been observed that the velocity of body waves that cross the focal region diminishes before a large earthquake. The time interval and amplitude of this decrease are somewhat related to the magnitude of the future earthquake. Sometimes the velocity tends to increase immediately before the occurrence. Although this phenomenon has been observed in some cases, in others changes in velocity were not followed by an earthquake. Thus, although this precursor raised great expectations, more controlled experiments have shown that it has little reliability.

Non-seismic phenomena of diverse types have been investigated, for example, changes in ground elevation or tilting, water levels in wells, strain in rocks, ground electric resistivity, electromagnetic fields, and emissions of radon gas. Observations of these physical parameters near an active fault are used to detect consistent patterns that may be correlated to the possible occurrence of an earthquake in the future. Geodetic and GPS measurements near faults can detect an accumulation of strain, but do not provide clues about its possible release. The presence of high-frequency electromagnetic signals, which is the basis of the proposed VAN method, also seems problematic (Varotsos and Lazaridou, 1991). New electromagnetic studies point to the precursor character of low-frequency emissions and effects in the ionosphere (Uyeda *et al.*, 2009; Hayakawa and Hobara, 2010).

Some changes in physical parameters that have been proposed as precursors have been justified theoretically in terms of the phenomenon called *dilatancy*, which implies an increase in volume and is observed in laboratory experiments on the fracture of rocks (Scholz *et al.*, 1973). According to the dilatancy of a material subject to increasing stress before the occurrence of a fracture, small cracks are produced that are then filled with fluids (water) present in the material and thus the volume increases. Owing to this phenomenon certain physical characteristics, such as the seismic wave velocity, seismicity, electric resistivity, and emission of radon gas, change. If the stress keeps on increasing, water is expelled and variations in some parameters (for example, the seismic velocity and seismicity) change sign. In this state, instabilities that determine its failure occur in the focal region. If this model were universally consistent, precursors would always have the same pattern in all cases, which is not the case.

Some modern programs of prediction have abandoned the idea of a unique precursor and search for a coherent pattern of convergence of several precursors. The rationale behind this practice is that, although no single precursor is a reliable predictor, the convergence of several of them may reveal consistent patterns. This approach was followed in the Parkfield (California) experiments in which many different types of measurements were made in order to detect precursor patterns (Lindh *et al.*, 1979). The present state of prediction research is very provisional and a solution in the near future cannot be foreseen (Evans, 1997). A question regarding complex non-linear dynamic systems has recently arisen from the new models of the occurrence of earthquakes. In these systems very small changes of initial or boundary conditions may produce very large effects on their subsequent development. Models of self-organized criticality have these properties and, if the occurrence of earthquakes follows these lines, it is possible that earthquakes may, in practice, not be predictable.

21.9 Summary

In this chapter we have applied quantitative knowledge about the generation and propagation of earthquake motion explained throughout the previous chapters to the problems presented in a descriptive form in [Chapter 2](#). We have seen the spatial distribution of earthquakes and its relation to the boundaries of tectonic plates where most earthquakes occur. The temporal distribution can be approximated by a Poisson distribution in which earthquakes are considered to be independent events save those separated by short intervals of distance and time. Main clusters of events are foreshocks and aftershocks, sequences, and swarms.

On a fault, the seismic cycle is formed by the accumulation of stress until it reaches a critical value and fracture is produced. However, the heterogeneity of the material does not allow periodic or regular cycles of earthquakes. Moreover, part of the accumulated stress may be released aseismically. Related to the seismic cycle is the concept of the characteristic earthquake, a concept that is also problematic. The relation of the number of earthquakes with magnitude follows the logarithmic Gutenberg and Richter law with a value of b near unity. Models for the occurrence of earthquakes, such as those of blocks and springs and cellular automata, have been discussed that model their fractal nature.

Earthquake spatial, time, and magnitude distribution and their source mechanisms provide information for the seismotectonics of a region. Their results confirm the tectonic processes related to plate motion, especially at plate boundaries. Finally we have presented a short account of the important practical problems of the assessment of seismic hazard and risk. The EEWS provides a possibility of giving an alert before the destructive waves arrive at a site. The possibility of earthquake prediction remains always a problem not to be forgotten.

Appendix 1 Vectors and tensors

A1.1 Definitions

Quantities representing physical entities can have one, two, three, or more components and thus can be defined by scalars, vectors, or tensors. In seismology, certain variables are scalars such as density, vectors such as displacements, and tensors such as stress and strain.

A scalar is a quantity with one component. Its value remains invariant under a change of the coordinate system.

A vector is a quantity with three components that transform under the rotation of the coordinate system as the coordinates of a point. To distinguish them from scalars, we represent them by bold face letters (**a**), or by a letter with a subindex (a_i) that takes the three values of the components in each coordinate system (a_1 , a_2 , and a_3). For the Cartesian coordinate system (x, y, z) or (x_1, x_2, x_3), a vector is represented by the three components or by their sum multiplied by the unit vectors in the direction of each axis (i, j, k) (Fig. A1.1):

$$\mathbf{a} = a_i = (a_x, a_y, a_z) = (a_1, a_2, a_3) = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}. \quad (\text{A1.1})$$

A vector modulus is a scalar given (in Cartesian coordinates) by

$$a = (a_1^2 + a_2^2 + a_3^2)^{1/2}. \quad (\text{A1.2})$$

If (x_1, x_2, x_3) is a Cartesian coordinate system and (x'_1, x'_2, x'_3) is another with the same origin and rotated by the angles between each pair of axes defined by

$$\cos(x_i, x'_j) = \beta_{ij}, \quad (\text{A1.3})$$

then the transformation of the components of a vector a_i referred to system x_i into a'_i , corresponding to the system x'_i , is given by

$$a'_i = \sum_{j=1}^3 \beta_{ij} a_j = \beta_{ij} a_j. \quad (\text{A1.4})$$

In the last term, we have used the convention that all repeated subindexes (in this case j) are summed up over their three values ($j = 1, 2, 3$) and it is not necessary to write the summation symbol.

A Cartesian tensor of second rank, represented by a bold face capital letter (**A**) or one with two subindexes (A_{ij}) can be defined, in a similar way to how we defined

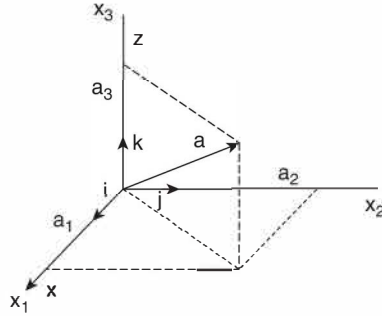


Fig. A1.1 Components of a vector in Cartesian coordinates.

a vector, as a quantity of nine components that transform under rotation of the coordinate axes in the form

$$B'_{ij} = \beta_{ik}\beta_{jl}B_{kl}. \quad (\text{A1.5})$$

An example of a second-rank tensor is the quantity β_{ij} formed by the cosines of the nine angles between the three axes of the two Cartesian coordinate systems x_i and x'_i . In this textbook, we generally use only vectors and Cartesian second-rank tensors. A more rigorous and general definition of tensors can be found in mathematical textbooks.

A1.2 Operations with vectors and tensors

Addition, subtraction, and multiplication by a scalar of vectors and tensors present no difficulty. In the subindex notation the same subindexes must be used throughout the whole expression:

$$\mathbf{c} = \mathbf{a} + \mathbf{b} - \mathbf{d}, \quad (\text{A1.6})$$

$$c_i = a_i + b_i - d_i, \quad (\text{A1.7})$$

$$\mathbf{d} = E\mathbf{a}, \quad (\text{A1.8})$$

$$d_i = Ea_i, \quad (\text{A1.9})$$

$$C_{ij} = A_{ij} + B_{ij} - D_{ij}. \quad (\text{A1.10})$$

The dot product of two vectors is a scalar given by

$$\mathbf{a} = \mathbf{b} \cdot \mathbf{c} = bc \cos(\mathbf{b}, \mathbf{c}), \quad (\text{A1.11})$$

where b and c are the moduli of the vectors $\mathbf{b}$ and $\mathbf{c}$. In the subindex notation, the dot product is

$$a = b_i c_i = b_1 c_1 + b_2 c_2 + b_3 c_3. \quad (\text{A1.12})$$

The dot product can also be expressed using the *Kronecker delta tensor* δ_{ij} , defined by

$$\delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}, \quad (\text{A1.13})$$

and is given by

$$a = \delta_{ij} b_i c_j. \quad (\text{A1.14})$$

The cross product of two vectors is a vector given by

$$\mathbf{a} = \mathbf{b} \times \mathbf{c} = bc \sin(\mathbf{b}, \mathbf{c}) \mathbf{e}, \quad (\text{A1.15})$$

where $\mathbf{e}$ is the unit vector normal to the plane that contains the vectors $\mathbf{b}$ and $\mathbf{c}$, which is positive in the direction of the right-hand rotation of $\mathbf{b}$ to $\mathbf{c}$.

The cross product of two vectors in subindex notation can be expressed by using the *alternating or permutation tensor* e_{ijk} :

$$e_{ijk} = \begin{cases} = 1, & i \neq j \neq k \text{ (even permutations)} \\ = -1, & i \neq j \neq k \text{ (odd permutations)} \\ = 0, & \text{(any index repeated)}. \end{cases} \quad (\text{A1.16})$$

The relation between the tensors e_{ijk} and δ_{ij} is

$$e_{ijk} e_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}. \quad (\text{A1.17})$$

The cross product is given by

$$\begin{aligned} a_i &= e_{ijk} b_j c_k, \\ a_i &= [(b_2 c_3 - b_3 c_2), (b_3 c_1 - b_1 c_3), (b_1 c_2 - b_2 c_1)]. \end{aligned} \quad (\text{A1.18})$$

We must remember that the cross product of two vectors is noncommutative and that the result is a pseudovector or axial vector. The scalar product of three vectors is

$$d = \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = e_{ijk} a_i b_j c_k. \quad (\text{A1.19})$$

Since all subindexes are repeated, they are summed and the result is a scalar. The triple cross product of three vectors is a vector, which can be expressed in terms of dot products as

$$\mathbf{d} = \mathbf{a} \times \mathbf{b} \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}, \quad (\text{A1.20})$$

$$d_i = e_{ijk} e_{klm} a_j b_l c_m = a_j c_j b_i - a_j b_j c_i. \quad (\text{A1.21})$$

The expression in subindex notation can be derived from equation (A1.17).

Two vectors can be used to construct a *second-rank tensor*, in what is called the *direct product* or *dyad*:

$$\mathbf{C} = \mathbf{a} \mathbf{b}, \quad (\text{A1.22})$$

$$C_{ij} = a_i b_j. \quad (\text{A1.23})$$

Because of this property, second-rank tensors are also called *diadics*. Vectors and second-rank tensors can be represented by matrices and their operation follows matrix algebra:

$$a_i = (a_1 \quad a_2 \quad a_3), \quad (\text{A1.24})$$

$$A_{ij} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}. \quad (\text{A1.25})$$

A1.3 Vector and tensor calculus

Scalars, vectors, and tensors that are functions of spatial coordinates are referred to as fields. Thus, in seismology, we use density, velocity, and stress fields. The most common operations in vector and tensor calculus, involving partial derivatives with respect to spatial coordinates, are *gradient*, *divergence*, and *curl*. In these cases, the symbolic vector operator ∇ , called *del* or *nabla*, is used. In Cartesian coordinates the nabla operator is defined as

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right) = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z}. \quad (\text{A1.26})$$

The gradient of a scalar is a vector:

$$\mathbf{a} = \nabla f = i \frac{\partial f}{\partial x} + j \frac{\partial f}{\partial y} + k \frac{\partial f}{\partial z}. \quad (\text{A1.27})$$

In subindex notation,

$$a_i = \frac{\partial f}{\partial x_i} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right) = f_{,i}. \quad (\text{A1.28})$$

In the last term, a comma followed by a subindex is used to indicate the partial derivatives with respect to the three spatial coordinates. Each derivative is the component of the resulting vector.

The gradient of a vector field f or f_i is a second-rank tensor:

$$\mathbf{B} = \nabla \mathbf{f}, \quad (\text{A1.29})$$

$$B_{ij} = \frac{\partial f_i}{\partial x_j} = f_{i,j}. \quad (\text{A1.30})$$

The divergence of a vector field is a scalar function:

$$g = \nabla \cdot \mathbf{f} = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z} = \frac{\partial f_i}{\partial x_i} = f_{i,i}. \quad (\text{A1.31})$$

The divergence of a tensor is a vector:

$$\begin{aligned}\mathbf{a} &= \nabla \cdot \mathbf{B}, \\ a_i &= \frac{\partial B_{ij}}{\partial x_j} = B_{ij,j}.\end{aligned}\tag{A1.32}$$

The divergence of the gradient of a scalar or vector field is the *Laplacian* (∇^2), an operator involving the second partial derivatives with respect to the three spatial coordinates. For Cartesian coordinates it is given by

$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = f_{,ii}.\tag{A1.33}$$

The Laplacian of a vector field is a vector formed by the Laplacian of their components:

$$\nabla^2 \mathbf{F} = \nabla^2 (F_1, F_2, F_3) = (\nabla^2 F_1, \nabla^2 F_2, \nabla^2 F_3).\tag{A1.34}$$

The curl of a vector function is another vector function. For Cartesian coordinates it is given by

$$\mathbf{d} = \nabla \times \mathbf{f} = \left(\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{\partial z}, \frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x}, \frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y} \right),\tag{A1.35}$$

$$d_i = e_{ijk} \frac{\partial f_k}{\partial x_j} = e_{ijk} f_{k,j}.\tag{A1.36}$$

The curl of the curl of a vector function is another vector function that can be expressed in terms of the gradient, divergence, and Laplacian:

$$\nabla \times \nabla \times \mathbf{f} = \nabla(\nabla \cdot \mathbf{f}) - \nabla^2 \mathbf{f}.\tag{A1.37}$$

This relation can be derived using index notation and equation (A1.17):

$$e_{ijk} e_{klm} f_{m,jl} = f_{j,ji} - f_{i,jj}.\tag{A1.38}$$

In equations involving only linear combinations of vectors and tensors, vector notation (bold face letters) is sufficient. In equations with calculus operations of vector and tensor functions, index notation is more convenient.

Appendix 2 Cylindrical and spherical coordinates

Cylindrical and spherical coordinates are used in many problems in seismology. The most basic equations are presented here in a brief form.

A2.1 Cylindrical coordinates

In problems that have symmetry with respect to an axis, for example, the propagation of surface waves on a flat Earth, cylindrical coordinates are commonly used. The position of a point in space in cylindrical coordinates is given by (ρ, ϕ, z) . The polar coordinates are a special case for a plane with only (ρ, ϕ) . With respect to the Cartesian coordinates (x, y, z) , ρ is the distance measured from the z axis in a plane parallel to (x, y) ($0 \leq \rho \leq \infty$), with the angle ϕ on the same plane as the x axis ($0 \leq \phi \leq 2\pi$) and distance z along the same axis ($-\infty \leq z < \infty$) (Fig. A2.1). The relations between cylindrical (or polar) and Cartesian coordinates are

$$x = \rho \cos \phi, \quad (\text{A2.1})$$

$$y = \rho \sin \phi. \quad (\text{A2.2})$$

The inverse transformations are

$$\rho = (x^2 + y^2)^{1/2}, \quad (\text{A2.3})$$

$$\phi = \tan^{-1}(y/x). \quad (\text{A2.4})$$

At each point of space of coordinates (ρ, ϕ, z) , we can establish a system of Cartesian coordinates in the direction of the unit vectors (e_ρ , e_ϕ , and e_z) with their positive senses defined by the positive increment of each coordinate (Fig. A2.1). These vectors change direction for each point of space and, thus, are functions of the coordinates. The relations between the unit vectors e_ρ , e_ϕ , and e_z and the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in the directions of the Cartesian coordinates $(x, y, \text{ and } z)$ are

$$e_\rho = \mathbf{i} \cos \phi + \mathbf{j} \sin \phi, \quad (\text{A2.5})$$

$$e_\phi = -\mathbf{i} \sin \phi + \mathbf{j} \cos \phi, \quad (\text{A2.6})$$

$$e_z = \mathbf{k}. \quad (\text{A2.7})$$

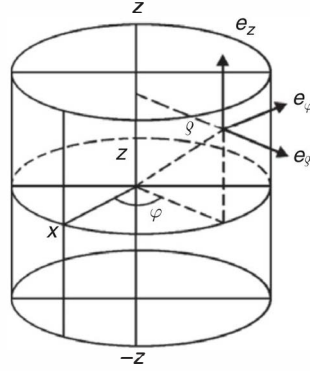


Fig. A2.1 Cylindrical coordinates. Unit vectors in the direction of coordinates.

The elements of distance ds , of the area dS over a cylindrical surface, and of volume dV are

$$ds^2 = d\rho^2 + \rho^2 d\phi^2 + dz^2, \quad (\text{A2.8})$$

$$dS = \rho d\phi dz, \quad (\text{A2.9})$$

$$dV = \rho d\rho d\phi dz. \quad (\text{A2.10})$$

Expressions for the gradient of a scalar function $f(\rho, \phi, z)$, and for the divergence and curl of a vector function $\mathbf{F}(\rho, \phi, z) = (F_\rho, F_\phi, F_z)$ are

$$\nabla f = \mathbf{e}_\rho \frac{\partial f}{\partial \rho} + \mathbf{e}_\phi \frac{1}{\rho} \frac{\partial f}{\partial \phi} + \mathbf{e}_z \frac{\partial f}{\partial z}, \quad (\text{A2.11})$$

$$\nabla \cdot \mathbf{F} = \frac{\partial F_\rho}{\partial \rho} + \frac{F_\rho}{\rho} + \frac{1}{\rho} \frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z}, \quad (\text{A2.12})$$

$$\begin{aligned} \nabla \times \mathbf{F} = & \mathbf{e}_\rho \left(\frac{1}{\rho} \frac{\partial F_z}{\partial \phi} - \frac{\partial F_\phi}{\partial z} \right) + \mathbf{e}_\phi \left(\frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial \rho} \right) \\ & + \mathbf{e}_z \left(\frac{\partial F_\phi}{\partial \rho} + \frac{F_\phi}{\rho} + \frac{1}{\rho} \frac{\partial F_\rho}{\partial \phi} \right). \end{aligned} \quad (\text{A2.13})$$

The Laplacian of a scalar function is

$$\nabla^2 f = \frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial f}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}. \quad (\text{A2.14})$$

Expressions in cylindrical coordinates are greatly simplified when there is symmetry with respect to ϕ . This type of symmetry with the principal axis in the z direction is called cylindrical symmetry.

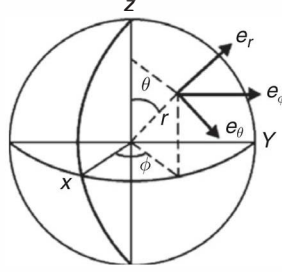


Fig. A2.2 Spherical coordinates. Unit vectors in the direction of coordinates.

A2.2 Spherical coordinates

The approximately spherical shape of the Earth makes the use of spherical coordinates very appropriate for problems of seismic wave propagation and normal modes in the Earth. In spherical coordinates, a point in space is given by $(r; \theta, \phi)$. With respect to the Cartesian coordinates (x, y, z) , r is the distance from the origin ($0 \leq r \leq \infty$), θ is the angle measured from the z axis ($0 \leq \theta \leq \pi$), and ϕ is the angle measured from the x axis to the projection of r onto the plane (x, y) ($0 \leq \phi \leq 2\pi$) (Fig. A2.2). The relations between the spherical and the Cartesian coordinates are

$$x = r \sin \theta \cos \phi, \quad (\text{A2.15})$$

$$y = r \sin \theta \sin \phi, \quad (\text{A2.16})$$

$$z = r \cos \theta. \quad (\text{A2.17})$$

The inverse relations are

$$r = (x^2 + y^2 + z^2)^{1/2}, \quad (\text{A2.18})$$

$$\theta = \cos^{-1} \left(\frac{z}{r} \right), \quad (\text{A2.19})$$

$$\phi = \tan^{-1} \left(\frac{y}{x} \right). \quad (\text{A2.20})$$

At each point of space $(r; \theta, \phi)$, we can establish a system of Cartesian coordinates in the directions of the unit vectors e_r , e_θ , and e_ϕ , with their positive senses in the directions of positive increments of the three coordinates (Fig. A2.2). These vectors change direction from point to point and are, thus, functions of the coordinates. The relations between the unit vectors e_r , e_θ , and e_ϕ and the unit vectors i , j , and k in the directions of the Cartesian coordinates (x, y, z) are

$$\mathbf{e}_r = \mathbf{i} \sin \theta \cos \phi + \mathbf{j} \sin \theta \sin \phi + \mathbf{k} \cos \theta, \quad (\text{A2.21})$$

$$\mathbf{e}_\theta = \mathbf{i} \cos \theta \cos \phi + \mathbf{j} \cos \theta \sin \phi + \mathbf{k} \sin \theta, \quad (\text{A2.22})$$

$$\mathbf{e}_\phi = -\mathbf{i} \sin \phi + \mathbf{j} \cos \phi. \quad (\text{A2.23})$$

The elements of distance ds , of the area dS on a spherical surface of constant radius, and of volume dV are

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2, \quad (\text{A2.24})$$

$$dS = r^2 \sin \theta d\theta d\phi, \quad (\text{A2.25})$$

$$dV = r^2 \sin \theta dr d\theta d\phi. \quad (\text{A2.26})$$

The gradient of a scalar function $f(r, \theta, \phi)$, and the divergence and curl of a vector function $\mathbf{F}(r, \theta, \phi) = (F_r, F_\theta, F_\phi)$, are

$$\nabla f = \mathbf{e}_r \frac{\partial f}{\partial r} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial f}{\partial \theta} + \mathbf{e}_\phi \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi}, \quad (\text{A2.27})$$

$$\nabla \cdot \mathbf{F} = \frac{1}{r} \frac{\partial}{\partial r} (r^2 F_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta F_\theta) + \frac{1}{r \sin \theta} \frac{\partial F_\phi}{\partial \phi}, \quad (\text{A2.28})$$

$$\begin{aligned} \nabla \times \mathbf{F} = & \mathbf{e}_r \frac{1}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (\sin \theta F_\phi) - \frac{\partial F_\theta}{\partial \phi} \right) - \mathbf{e}_\theta \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial F_r}{\partial \phi} - \frac{\partial}{\partial r} (r \sin \theta F_\phi) \right) \\ & + \mathbf{e}_\phi \frac{1}{r} \left(\frac{\partial}{\partial r} (r F_\theta) - \frac{\partial F_r}{\partial \theta} \right) \end{aligned} \quad (\text{A2.29})$$

The Laplacian of a scalar function is

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}. \quad (\text{A2.30})$$

Expressions in spherical coordinates are simplified when there is some sort of symmetry. Symmetry with respect to ϕ and θ is called spherical symmetry and there is only a dependence on r . In many problems symmetry with respect to ϕ is used and there are only dependences on r and θ .

Appendix 3 Bessel and Legendre functions

A3.1 Bessel functions

In many problems with cylindrical symmetry, solutions of the Laplace, wave, and diffusion differential equations are given in terms of *Bessel functions*. This is the case in seismology with the solutions of wave equations in cylindrical and spherical coordinates. Bessel functions are solutions of *Bessel differential equations*:

$$\frac{d^2 f}{dx^2} + \frac{1}{x} \frac{df}{dx} + \left(1 - \frac{n^2}{x^2}\right) f = 0. \quad (\text{A3.1})$$

Solutions with finite values for $x = 0$ are Bessel functions of the first class $J_n(x)$, with integer values of n . These functions are expressed as infinite sums of powers of x :

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(m+n)!} \left(\frac{x}{2}\right)^{n+2m}. \quad (\text{A3.2})$$

For $n = 0$,

$$J_0(x) = 1 + \frac{x^2}{2^2 1!} + \frac{x^4}{2^4 2!} + \frac{x^6}{2^6 3!} + \dots \quad (\text{A3.3})$$

For other values of $n > 0$, they can be obtained from those of order zero:

$$J_n(x) = (-x)^n \left(\frac{1}{x} \frac{d}{dx} \right)^n J_0(x). \quad (\text{A3.4})$$

Several recurrence formulas relate Bessel functions of different orders and their derivatives; among them are

$$J'_n(x) = \frac{n}{x} J_n(x) - J_{n+1}(x), \quad (\text{A3.5})$$

$$2J'_n(x) = J_{n-1}(x) - J_{n+1}(x). \quad (\text{A3.6})$$

Solutions of differential Bessel equations with non-finite values for $x = 0$ are Bessel functions of the second kind $Y_n(x)$. They can be deduced from those of the first kind:

$$Y_n(x) = -\frac{\cos(n\pi)J_n(x) - J(x)}{\sin(n\pi)}. \quad (\text{A3.7})$$

From Bessel functions of the first and second kinds, we form Bessel functions of the third kind, or *Hankel functions*:

$$H_n^1(x) = J_n(x) + iY_n(x), \quad (\text{A3.8})$$

$$H_n^2(x) = J_n(x) - iY_n(x). \quad (\text{A3.9})$$

For large values of x ($|x| \gg 1$ and $|x| \gg n$), we have for Bessel functions of the first and second classes the following asymptotic expressions:

$$J_n(x) \simeq \left(\frac{2}{\pi x}\right)^{1/2} \cos\left(x - \frac{\pi}{4} - \frac{n\pi}{2}\right), \quad (\text{A3.10})$$

$$Y_n(x) \simeq \left(\frac{2}{\pi x}\right)^{1/2} \sin\left(x - \frac{\pi}{4} - \frac{n\pi}{2}\right). \quad (\text{A3.11})$$

A3.2 Spherical Bessel functions

Spherical Bessel functions of the first class are derived from Bessel functions of the first class of non-integer order:

$$j_n(x) = \left(\frac{\pi}{2x}\right)^{1/2} J_{n+1/2}(x). \quad (\text{A3.12})$$

These functions are solutions of the differential equation

$$\frac{d^2 f}{dx^2} + \frac{2}{x} \frac{df}{dx} + \left(1 - \frac{n(n+1)}{x^2}\right) f = 0. \quad (\text{A3.13})$$

This differential equation appears in the resolution of the wave equation in spherical coordinates for the dependence on r . Their name comes from their relation to trigonometric functions:

$$j_n(x) = (-x)^n \left(\frac{1}{x} \frac{d}{dx}\right)^n \frac{\sin x}{x}. \quad (\text{A3.14})$$

Complex solutions are given in terms of *Hankel spherical functions*:

$$h_n(x) = j_n(x) + in_n(x), \quad (\text{A3.15})$$

where

$$n_n(x) = (-1)^{n+1} \left(\frac{\pi}{2x}\right)^{1/2} J_{-n-1/2}(x).$$

Functions $j_n(x)$ and $h_n(x)$ have asymptotic approximations ($x \rightarrow \infty$), given by

$$j_n(x) \simeq \frac{1}{x} \cos \left[x - \frac{\pi}{2}(n+1) \right], \quad (\text{A3.16})$$

$$h_n(x) \simeq \frac{1}{x} \exp i \left[x - \frac{\pi}{2}(n+1) \right]. \quad (\text{A3.17})$$

A3.3 Legendre functions

Solutions for many problems in spherical coordinates, especially those involving Laplace and wave equations, are expressed in terms of *Legendre functions*. We start with the *Legendre differential equation*,

$$(1-x^2) \frac{d^2 f}{dx^2} - 2x \frac{df}{dx} + n(n+1) = 0. \quad (\text{A3.18})$$

Solutions for $|x| \leq 1$ can be expressed in terms of Legendre functions of the first and second kind, $P_n(x)$ and $Q_n(x)$. Those of the first kind for integer values of n , called *Legendre polynomials*, are given by

$$P_n(x) = \sum_{s=0}^m \frac{(-1)^s (2n-2s)! x^{n-2s}}{s!(n-s)!(n-2s)! 2^n}, \quad (\text{A3.19})$$

where $m = n/2$ or $(n-1)/2$, so that it is always an integer. Legendre polynomials can be generated from certain functions, such as

$$(1-2xz+z^2)^{-1/2} = \sum_{n=0}^{\infty} z^n P_n(x), \quad (\text{A3.20})$$

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2-1)^n. \quad (\text{A3.21})$$

Some recurrence formulas for Legendre polynomials of different order and their derivatives are

$$P'_{n+1}(x) - xP'_n(x) = (n+1)P_n(x), \quad (\text{A3.22})$$

$$(2n+1)xP_n(x) = (n+1)P_{n+1}(x) + nP_{n-1}(x). \quad (\text{A3.23})$$

Legendre polynomials are orthogonal, so they satisfy

$$\int_{-1}^{+1} P_m(x) P_n(x) dx = \delta_{nm} \frac{2}{2n+1}. \quad (\text{A3.24})$$

The first polynomials are

$$P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x), P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3).$$

A3.4 Associate Legendre functions

A general form of the differential Legendre equation is

$$(1 - x^2) \frac{d^2 f}{dx^2} - 2x \frac{df}{dx} + \left(n(n+1) - \frac{m^2}{1-x^2} \right) f = 0. \quad (\text{A3.25})$$

This equation reduces to (A3.18) for $m = 0$. The solutions of equation (A3.25) are *associate Legendre functions* of order n and grade m . They can be obtained from Legendre polynomials via

$$P_n^m(x) = (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_n(x). \quad (\text{A3.26})$$

Some recurrence formulas are

$$(n - m) P_n^m(x) = (2n - 1)x P_{n-1}^m(x) - (n + m - 1) P_{n-2}^m(x), \quad (\text{A3.27})$$

$$(n - m) P_n^m(x) = nx P_{n-1}^m(x) - (1 - x^2) \frac{d}{dx} P_{n-1}^m(x). \quad (\text{A3.28})$$

The condition of orthogonality is

$$\int_{-1}^{+1} P_n^m(x) P_l^m(x) dx = \delta_{nl} \frac{2(n+m)!}{(2n+1)(n-m)!}, \quad (\text{A3.29})$$

$$\int_{-1}^{+1} P_n^m(x) P_m^l(x) dx = \delta_{ml} \frac{(n+m)!}{m(n+m)!}. \quad (\text{A3.30})$$

For $m = 0$, associate Legendre functions are equivalent to Legendre polynomials. For $m > 0$, the first associate functions are

$$P_1^1(x) = (1 - x^2)^{1/2}, \quad P_2^1(x) = 3x(1 - x^2)^{1/2}$$

$$P_2^2(x) = 3(1 - x^2), \quad P_3^1(x) = \frac{1}{2}(1 - x^2)^{1/2}(15x^2 - 3).$$

In spherical coordinates, solutions of Laplace and wave equations using the method of separation of variables result in equation (A3.25) with $x = \cos \theta$, and a harmonic equation

for ϕ . Solutions can be given in terms of the products of associate Legendre and trigonometric functions:

$$S_n^m(\theta, \phi) = [A_n^m \cos(m\phi) + B_n^m \sin(m\phi)] P_n^m(\cos \theta). \quad (\text{A3.31})$$

These functions are called *spherical surface harmonics* or *tesseral harmonics*. If the problem has symmetry with respect to ϕ , the resulting equation is (A3.18) with $x = \cos \theta$ and the solution is given in terms of Legendre polynomials, which, in this case, are also called *zonal harmonics*. If we introduce into (A3.31) a factor that depends on positive or negative powers of r , we obtain the *solid harmonics*.

Appendix 4 Fourier and Laplace transforms

Seismological data representing ground displacements, velocities, and accelerations produced by earthquakes and recorded at a point on the Earth's surface are functions of time. Powerful tools for their analysis are the *Fourier* and *Laplace transform* theorems which allow one to change from the time to the frequency domain and vice versa. The fundamental ideas of these important mathematical tools are presented in a very brief form. Although the Fourier and Laplace theorems can be applied to very general functions of any variable, here we will treat only time functions (Bracewell, 1986).

A4.1 Periodic functions

A periodic function $f(t)$ of period T and frequency $\omega = 2\pi/T$ that satisfies very general conditions of continuity, such as being absolutely integrable, can be expressed, according to Fourier's theorem, as an infinite series of harmonic functions:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{2\pi n}{T} t\right) + b_n \sin\left(\frac{2\pi n}{T} t\right) \right]. \quad (\text{A4.1})$$

The coefficients a_n and b_n are obtained in the forms

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos\left(\frac{2\pi n}{T} t\right) dt, \quad (\text{A4.2})$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin\left(\frac{2\pi n}{T} t\right) dt. \quad (\text{A4.3})$$

If we make the substitution $\omega_n = 2\pi n/T$, then the arguments of the sine and cosine functions are given in terms of the frequencies $\omega_n = n\omega$, which are integer multiples of the frequency of the function $f(t)$. Equation (A4.1) can be expressed in exponential form as

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{i\omega_n t}, \quad \omega = \frac{2\pi}{T}. \quad (\text{A4.4})$$

The coefficients F_n have complex values:

$$F_n = \frac{1}{2}(a_n + ib_n), \quad (\text{A4.5})$$

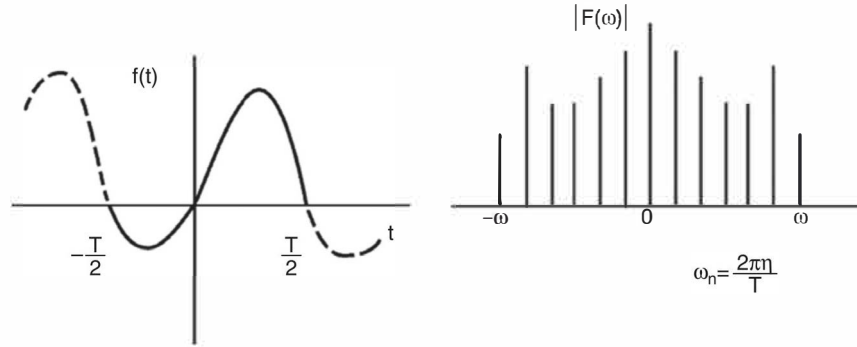


Fig. A4.1 A periodic continuous function and its discrete transform.

$$F_{-n} = \frac{1}{2}(a_n - ib_n). \quad (\text{A4.6})$$

These coefficients are found by writing

$$F_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i\omega_n t} dt. \quad (\text{A4.7})$$

The complex coefficients F_n can be expressed as

$$F_n = |F_n| e^{i\phi_n} = A_n e^{i\phi_n}, \quad (\text{A4.8})$$

$$A = \frac{1}{2}(a_n^2 + b_n^2)^{1/2}, \quad (\text{A4.9})$$

$$\phi_n = \tan^{-1}(b_n + a_n). \quad (\text{A4.10})$$

Thus, F_n is the complex Fourier transform of $f(t)$ and, since this is a reciprocal relation, $f(t)$ is the inverse transform of F_n . The values $A_n = |F_n|$ constitute the amplitude spectrum and those of ϕ_n constitute the phase spectrum. The square of the amplitude spectrum is called the power spectrum $P_n = |F_n|^2$. Thus, a periodic continuous function $f(t)$ has a discrete transform with values only for frequencies $\omega_n = n\omega$, which are integer multiples (positive or negative) of its own frequency (Fig. A4.1).

A4.2 Non-periodic functions

A non-periodic function $f(t)$ defined for all values of t and that satisfies the already mentioned continuity conditions can be expressed using the Fourier integral as

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega, \quad (\text{A4.11})$$

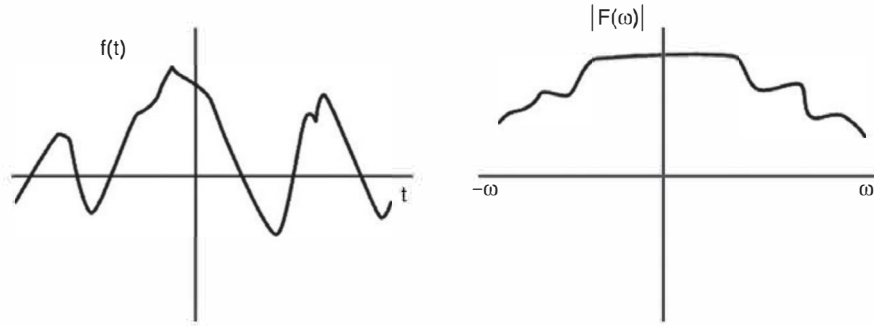


Fig. A4.2 A non-periodic infinite continuous function and its infinite continuous transform.

where $F(\omega)$, the complex Fourier transform, is given by

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt. \quad (\text{A4.12})$$

Thus, upon substitution of (A4.12) into (A4.11), the Fourier integral theorem is given by

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt d\omega. \quad (\text{A4.13})$$

The Fourier transform $F(\omega)$ is a continuous function of the frequency ω that is defined over the whole range of positive and negative frequencies. Although $f(t)$ can be a real function, $F(\omega)$ is a complex function with real and imaginary parts, $F(\omega) = R(\omega) + iI(\omega)$, that may be also expressed in the form

$$F(\omega) = A(\omega) e^{i\phi(\omega)}, \quad (\text{A4.14})$$

$$A(\omega) = |F(\omega)| = [R(\omega)^2 + I(\omega)^2]^{1/2}, \quad (\text{A4.15})$$

$$\phi(\omega) = \tan^{-1} [I(\omega)/R(\omega)], \quad (\text{A4.16})$$

where $A(\omega)$ is the *amplitude spectrum* and $\phi(\omega)$ is the *phase spectrum*, both of which are continuous functions of the frequency (Fig. A4.2).

A4.3 Convolution and correlation

An important operation in the analysis of temporal functions such as seismological data is the *convolution* of two functions, which is given by

$$h(\tau) = \int_{-\infty}^{\infty} f(t) g(\tau - t) dt = f(t) * g(t). \quad (\text{A4.17})$$

The convolution function $h(\tau)$ is a function of the time shift τ . In the shifted function $g(\tau - t)$, the independent variable (t) has changed sign. A property of convolution is that its transform is the product of the transforms of the two functions:

$$H(\omega) = F(\omega)G(\omega). \quad (\text{A4.18})$$

On passing from the time to the frequency domain, convolution in time changes into a product in frequency. Thus, it may be advantageous instead of doing time convolutions to transform into the frequency domain to take products and then change back into the time domain by means of the inverse transform.

Another operation that can be applied to two functions is their *correlation*:

$$q(\tau) = \int_{-\infty}^{\infty} f(t)g(t + \tau) dt. \quad (\text{A4.19})$$

The function $q(\tau)$ is the correlation function of $f(t)$ and $g(t)$. If the functions are different it is called *cross correlation*, whereas if the operation is performed on the same function the result is the autocorrelation function,

$$p(\tau) = \int_{-\infty}^{\infty} f(t)f(t + \tau) dt. \quad (\text{A4.20})$$

The transform of the *autocorrelation* function $p(\tau)$ is the *power spectrum* of the original function $f(t)$,

$$P(\omega) = |F(\omega)|^2 = A^2(\omega). \quad (\text{A4.21})$$

When we calculate the autocorrelation function, we lose information about the phases of its transform. Autocorrelation functions are always even functions and their transforms have only real parts.

A4.4 Sampled functions of finite duration

The digital form of modern seismologic data and the use of digital computers for their analysis determines that time functions representing them are sampled at certain time increments and have finite duration. Both conditions have consequences for their Fourier analysis.

If a continuous function $f(t)$ is defined only in the interval $(-t_m, t_m)$ and we do not know what values it takes outside this interval, we cannot calculate its transform, since this requires that the function be known for all times. To resolve this difficulty, we must make a hypothesis about the values of the function outside the known range; either that it is zero or that it repeats as a periodic function of period $2t_m$. The first hypothesis is equivalent to multiplying the function $f(t)$ by a rectangular function of amplitude unity and length $2t_m$ (a rectangular window). Since multiplication in one domain corresponds to a convolution in the other, the spectrum of $f(t)$ is convoluted with the transform of the rectangular window. In the second hypothesis, $f(t)$ becomes a periodic function and, according to (A4.4), can be represented by an infinite series whose coefficients are

$$F_n = \frac{1}{2t_m} \int_{-t_m}^{t_m} f(t) \exp\left(-\frac{i\pi n t}{t_m}\right) dt. \quad (\text{A4.22})$$

Although $f(t)$ is a continuous function, since it is considered to be periodic, its transform is discrete with values at increments in sampling frequency of $\Delta\omega = \pi/t_m$:

$$F_n = F\left(\frac{\pi n}{t_m}\right). \quad (\text{A4.23})$$

A sampled function is one that is known only for discrete values of the variable. If the sampling increment Δt is constant, then

$$f_n = f(n \Delta t), \quad n = 0, 1, 2, 3, \dots \quad (\text{A4.24})$$

According to the relation between a function and its transform, if f_n is a discrete function at increments Δt , then, using (A4.4) and (A4.7), it can be considered as the inverse transform of a periodic function $F(\omega)$ with period $2\pi/\Delta t$ (A4.4). Therefore, the transform function $F(\omega)$ is limited to a band of frequencies between $-\omega_N$ and ω_N , where $\omega_N = \pi/\Delta t$ is called the *Nyquist frequency* (for $\sigma = \omega/(2\pi)$, $\sigma_N = 1/(2\Delta t)$). Frequencies higher than the Nyquist frequency are folded together in the phenomenon known as *aliasing*. In this way f_n is given in terms of its transform by

$$f_n = \frac{1}{2\omega_N} \int_{-\omega_N}^{\omega_N} F(\omega) \exp\left(\frac{i\pi n \omega}{2\omega_N}\right) d\omega. \quad (\text{A4.25})$$

If the number of samples is finite, say equal to $2N$, f_n is a sampled function of finite duration, $2t_m = 2N \Delta t$. In practice, all of the functions we use in seismological analysis are of this type. According to (A4.22), (A4.23), and (A4.25), its transform is discrete and is limited to a band of frequencies:

$$F_n = F\left(\frac{n\pi}{N \Delta t}\right), \quad n = -N, \dots, -2, -1, 0, 1, 2, \dots, N. \quad (\text{A4.26})$$

Thus, the transform is limited to the frequency band $-\pi/\Delta t \leq \omega_n \leq \pi/\Delta t$ and the Nyquist frequency is $\omega_N = \pi/\Delta t$ ($\sigma_N = 1/(2\Delta t)$). The relation between f_n and F_n is given by the pair of transform equations

$$f_n = \frac{1}{N} \sum_{k=-N}^N F_k \exp\left(\frac{i\pi n k}{N}\right), \quad n = -N, \dots, N, \quad (\text{A4.27})$$

$$F_n = \sum_{k=-N}^N f_k \exp\left(\frac{i\pi n k}{N}\right), \quad n = -N, \dots, N. \quad (\text{A4.28})$$

These relations are known as the discrete Fourier transforms for sampled functions of limited length. The time function f_n is defined in the interval $-N \Delta t \leq t \leq N \Delta t$ and sampled at increments Δt . Its transform F_n is defined in the frequency band $-\pi/\Delta t \leq \omega \leq \pi/\Delta t$ and sampled at increments $\Delta\omega = \pi/(N \Delta t)$. If f_n is real, $F_{-n} = F_n$, and the transform is defined by the N values of positive frequencies (Fig. A4.3). Therefore, if we want to increase the frequency band of the spectrum, we must decrease the value of the

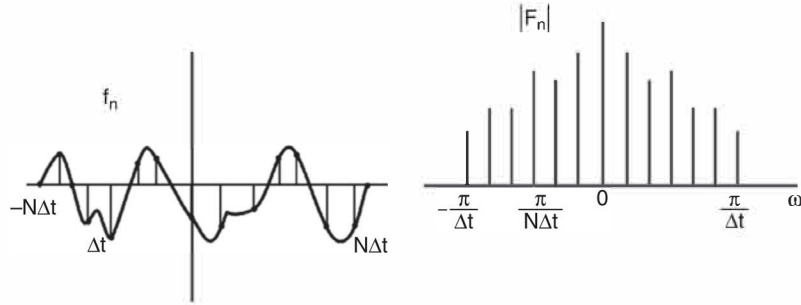


Fig. A4.3

A non-periodic sampled function of limited length and its discrete limited transform.

increment in sampling time, and if we want to increase its resolution (smaller frequency increments), we must increase the length of the function.

A4.5 Laplace transform

The *Laplace transform* of a function $f(t)$ is defined as (Bracewell, 1986)

$$L[f(t)] = F(s) = \int_{-\infty}^{+\infty} f(t)e^{-st} dt, \quad (\text{A4.29})$$

where $s = c + i\omega$ is a complex variable. The main difference between the Fourier $F(\omega)$, and Laplace transform, $F(s)$, is that the $F(\omega)$ is a function, that may be complex, of a real variable (ω), while $F(s)$ is a function of a complex variable (s).

The inverse Laplace transform is defined as

$$f(t) = \frac{1}{2\pi i} \int_{c-i\omega}^{c+i\omega} F(s)e^{st} ds. \quad (\text{A4.30})$$

The Laplace transform is a linear transform with many properties similar to the Fourier transform. An important property for us is that the derivative in the time domain corresponds to multiplying by s in the Laplace domain,

$$\frac{df(t)}{dt} = sF(s). \quad (\text{A4.31})$$

Estimation of the inverse Laplace transform is a problem not always easy to solve; for this reason it is convenient to decompose $F(s)$ in several single functions with a well-known inverse Laplace transform using the property

$$L[f_1(t) + f_2(t)] = F_1(s) + F_2(s). \quad (\text{A4.32})$$

If we have the function $F(s) = A(s)/B(s)$, where $A(s)$ and $B(s)$ are polynomials of s ,

$$F(s) = \frac{A(s)}{k_0(s - a)(s - b)(s - c) \dots}, \quad (\text{A4.33})$$

where k_0 is the coefficient of the larger s power of $B(s)$. We can write this equation as a sum of fractions,

$$F(s) = \frac{k_1}{s - a} + \frac{k_2}{s - b} + \frac{k_3}{s - c} + \dots, \quad (\text{A4.34})$$

where $a, b, c, \dots$ are the roots of $B(s)$. These roots may be real or complex. Now it is very easy to obtain the inverse Laplace transform of these fractions using the pairs of Laplace transforms:

$$F(s) = \frac{1}{1 + Ts} \Leftrightarrow f(t) = \frac{1}{T} e^{-t/T}. \quad (\text{A4.35})$$

The convolution theorem for the Laplace transform is similar to that of the Fourier transform. If $f(t)$ and $g(t)$ are functions in the time domain where $F(s)$ and $G(s)$ are their Laplace transforms, the Laplace transform of their convolution is the product of their Laplace transforms:

$$L(f * g) = F(s) G(s). \quad (\text{A4.36})$$

Appendix 5 Parameters of the Earth

Equatorial radius	6 378 136 m
Polar radius	6 356 750 m
Radius of the sphere of equal volume	6 370 800 m
Volume	$1.083 \times 10^{21} \text{ m}^3$
Mass	$5.973 \times 10^{24} \text{ kg}$
Mean density	5.515 g cm^{-3}
Surface area	$5.1 \times 10^{14} \text{ m}^2$
Polar moment of inertia	$8.0378 \times 10^{37} \text{ kg m}^2$
Equatorial moment of inertia	$8.0115 \times 10^{37} \text{ kg m}^2$
Flattening	3.3528×10^{-3}
Angular velocity	$7.2921 \times 10^{-5} \text{ s}^{-1}$
Core radius	3486 km
Inner core radius	1271 km
Core mass	$1.883 \times 10^{24} \text{ kg}$
Thickness of mantle	2890 km
Mantle mass	$4.09 \times 10^{24} \text{ kg}$
Mean distance to the Sun	$1.496 \times 10^8 \text{ km}$
Mean distance to the Moon	$3.844 \times 10^5 \text{ km}$

Appendix 6 The interior of The Earth

The interior of the Earth	Depth (km)	Density (g cm ⁻³)	Acceleration due to gravity (ms ⁻²)	Pressure (10 ⁴ MPa)	Bulk (K) (10 ⁴ MPa)	Shear (μ) (10 ⁴ MPa)	P wave velocity (α) (km s ⁻¹)	S wave velocity (β) (km s ⁻¹)	Temperature (K)
Crust	10–30	2.9	9.83	0.06	6.82	4.41	6.57	3.82	280
Upper mantle	40	3.33	9.84	0.12	13.19	5.18	8.12	4.42	650
	220	3.36	9.90	0.71	15.29	7.41	8.06	4.35	1610
	400	3.72	9.97	1.34	18.99	9.06	9.13	5.22	1910
Lower mantle	670	4.38	10.01	2.38	29.26	15.48	10.75	5.95	2375
	1200	4.69	9.94	4.78	38.50	19.96	11.78	6.52	2525
	2885	5.57	10.68	13.58	65.56	29.38	13.72	7.26	2934
Outer core	2890	9.90	10.68	13.58	64.41	0.0	8.06	0.0	3160
	3800	11.11	8.42	22.75	96.33	0.0	9.31	0.0	3650
	5150	12.17	4.40	32.89	130.47	0.0	10.36	0.0	4170
Inner core	5155	12.76	4.40	32.89	134.34	15.67	11.03	3.50	4190
	6371	13.09	0.0	36.39	142.53	17.61	11.26	3.67	4290

Appendix 7 Important earthquakes

Date	Region	Casualties	Magnitude
21 07 365	Aegean		
575	Antioch, Turkey	10 000	
856	Corinth, Greece	10 000	
27 09 1290	Chilhi, China	100 000	
02 02 1428	Olot, Spain	800	
05 12 1456	Campania, Italy	30 000	
23 01 1556	Shaanxi, China	300 000	
04 02 1729	Quito, Ecuador	40 000	
01 11 1755	Lisbon, Portugal	20 000	
04 02 1783	Calabria, Italy	50 000	
16 08 1868	Peru–Ecuador	100 000	
16 12 1875	Naples, Italy	12 000	
25 12 1884	Andalusia, Spain	749	
31 08 1886	Charleston, USA	60	
28 10 1891	Mino–Owari, Japan	7000	
15 06 1896	Riku–Ugo, Japan	22 000	
12 06 1897	Assam, India	1500	8.7
18 04 1906	San Francisco, USA	700	8.2
28 12 1908	Messina, Italy	120 000	7.5
16 12 1920	Shaanxi, China	30 000	8.5
01 09 1923	Kwanto, Japan	99 300	8.2
15 01 1934	Nepal, India	10 700	8.3
24 01 1939	Chillan, Chile	30 000	7.7
26 12 1939	Erzincan, Turkey	23 000	8.0
15 08 1950	Assam, India	574	8.6
04 11 1952	Kamchatka, USSR	600	8.5
22 05 1960	Valdivia, Chile	5700	8.5
28 03 1964	Anchorage, USA	131	8.5
31 08 1968	Dashy–Bayaz, Iran	11 600	7.4
31 05 1970	Ancash, Peru	68 000	7.8
23 12 1972	Managua, Nicaragua	5000	6.2
04 02 1976	Guatemala	22 000	7.9
27 07 1976	Tangshan, China	650 000	7.6
04 03 1977	Vrancea, Romania	2000	7.2

Date	Region	Casualties	Magnitude
16 09 1978	Iran	15 000	7.4
19 09 1985	Michoacan, Mexico	20 000	8.1
07 12 1988	Armenia	25 000	7.0
18 10 1989	Loma Prieta, USA	62	7.0
21 06 1990	Iran	50 000	7.7
29 09 1993	Latur, India	11 000	6.3
16 01 1995	Kobe, Japan	5502	7.5
17 08 1999	Izmit, Turkey	17 000	7.6
21 09 1999	Jiji, Taiwan	2415	7.7
26 01 2001	Gujarat, India	20 085	7.7
23 06 2001	Arequipa, Peru	102	8.4
26 12 2003	Bam, Iran	30 000	6.6
26 12 2004	Sumatra, Indonesia	228 000	9.1
08 10 2005	Kashmir, Pakistan	95 000	7.6
12 05 2008	Sichuan, China	69 197	7.9
12 01 2010	Port-au-Prince, Haiti	316 000	8.8
27 02 2010	Maule, Chile	577	9.0
11 03 2011	Tôhoku, Japan	18 493	9.0
25 04 2015	Nepal	9018	7.8

Bibliography

General seismology textbooks

- Aki, K. and P. G. Richards (1980). *Quantitative Seismology. Theory and Methods* (2 volumes). W. H. Freeman, San Francisco. 932 pp.
- Båth, M. (1973). *Introduction to Seismology*. John Wiley, New York; Birkhäuser, Basel. 395 pp.
- Ben Menahem, A. and S. J. Singh (1981). *Seismic Waves and Sources* (2 volumes). Springer, Berlin. 1168 pp.
- Bolt, B. A. (1978). *Earthquakes, A Primer*. W. H. Freeman, San Francisco. 241 pp.
- Bullen, K. E. (1947). *An Introduction to the Theory of Seismology*. Cambridge University Press, Cambridge. 296 pp.
- Bullen, K. E. and B. A. Bolt (1985). *An Introduction to the Theory of Seismology*. Cambridge University Press, Cambridge. 499 pp.
- Byerly, P. (1942). *Seismology*. Prentice Hall, New York. 256 pp.
- Dahlen, F. A. and J. Tromp (1998). *Theoretical Global Seismology*. Princeton University Press, Princeton. 1025 pp.
- Doyle, H. (1995). *Seismology*. John Wiley, New York. 218 pp.
- Galitzin, B. (1914). *Vorlesungen der Seismometrie*. Teubner, Leipzig. 560 pp.
- Gershanik, S. (1995). *Sismología*. Universidad Nacional de la Plata, Buenos Aires. 826 pp.
- Gubbins, D. (1990). *Seismology and Plate Tectonics*. Cambridge University Press, Cambridge. 339 pp.
- Hobbs, H. (1907). *Earthquakes. An Introduction to Seismic Geology*. Appleton, New York. 336 pp.
- Hoernes, R. (1893). *Erdbebenkunde*. Veit, Leipzig. 452 pp.
- Lay, T. and T. C. Wallace (1995). *Modern Global Seismology*. Academic Press, San Diego. 521 pp.
- Macelwane, J. B. and F. W. Sohon (1936). *Introduction to Theoretical Seismology. Part I, Geodynamics, and Part II, Seismometry*. John Wiley, New York. 366 pp. and 149 pp.
- Madariaga, R. and G. Perrier (1991). *Tremblement de terre*. Presses CNRS, Paris. 210 pp.
- Milne, J. (1898). *Seismology*. Kegan, Trench and Trübner. London. 320 pp.
- Montessus de Ballore, F. (1911). *La Sismologie Moderne*. A. Colin, Paris. 284 pp.
- Pujol, J. (2003). *Elastic Wave Propagation and Generation*. Cambridge University Press, Cambridge. 444 pp.
- Richter, C. F. (1958). *Elementary Seismology*. W. H. Freeman, San Francisco. 768 pp.
- Sawarensky, E. F. and D. P. Kirnos (1960). *Elemente der Seismologie und Seismometrie*. Akademie Verlag, Berlin. 512 pp.
- Shearer, P. M. (1999). *Introduction to Seismology*. Cambridge University Press, Cambridge. 260 pp.
- Sieberg, A. (1904). *Handbuch der Erdbebenkunde*. Braunschweig and Bieweg. Wiesbaden. 327 pp.

- Stein, S. and M. Wysession (2003). *An Introduction to Seismology. Earthquakes and Earth Structure*. Blackwell, Malden. 512 pp.
- Udias, A. (1999). *Principles of Seismology*. Cambridge University Press, Cambridge. 475 pp.
- Udias, A. and J. Mezcua (1996). *Fundamentos de sismología*. UCA Editores, San Salvador. 200 pp.

Special topics in seismology

- Babuska, V. and M. Cara (1991). *Seismic Anisotropy in the Earth*. Kluwer Academic, Dordrecht. 217 pp.
- Båth, M. (1968). *Mathematical Aspects of Seismology*. Developments in Solid Earth Geophysics 4. Elsevier, Amsterdam. 414 pp.
- Bolt, B. A. (1982). *Inside the Earth. Evidence from Earthquakes*. W. H. Freeman, San Francisco. 191 pp.
- Boschi, E., G. Ekström and A. Morelli, eds. (1996). *Seismic Modelling of Earth Structure*. Editrice Compositori, Bologna. 572 pp.
- Červený, V. (2001). *Seismic Ray Theory*. Cambridge University Press, Cambridge. 713 pp.
- Chapman, C. H. (2004). *Fundamentals of Seismic Wave Propagation*. Cambridge University Press, Cambridge. 634 pp.
- Claerbout, J. F. (1985). *Imaging the Earth Interior*. Blackwell Science Press, Oxford. 398 pp.
- Coburn, A. and R. Spence (1992). *Earthquake Protection*. John Wiley and Sons, Chichester. 355 pp.
- Davison, C. (1927). *The Founders of Seismology*. Cambridge University Press, Cambridge. 240 pp.
- Dziewonski, A. M. and E. Boschi, eds. (1980). *Physics of the Earth's Interior*. North-Holland, Amsterdam. 715 pp.
- Ewing, W. M., W. S. Jardetzky, and F. Press (1957). *Elastic Waves in Layered Media*. McGraw-Hill, New York. 380 pp.
- Fukuyama, E., ed. (2009). *Fault-Zone Properties and Earthquake Rupture Dynamics*. Elsevier, Amsterdam, 308 pp.
- Gibowicz, S. J. and A. Kijko (1994). *An Introduction to Mining Seismology*. Academic Press, San Diego. 399 pp.
- Gutenberg, B. and C. F. Richter (1949, 1954). *Seismicity of the Earth and Associated Phenomena*. Princeton University Press, Princeton. 273 pp.
- Iyer, H. N. and K. Hirahara, eds. (1993). *Seismic Tomography. Theory and Practice*. Chapman and Hall, London. 849 pp.
- James, D. E., ed. (1989) *The Encyclopedia of Solid Earth Geophysics*. Van Nostrand-Reinhold, New York (entries on seismological subjects).
- Jeffreys, H. (1924, 1927, 1959). *The Earth, its Origin, History and Physical Constitution*. Cambridge University Press, Cambridge. 278 pp.
- Kanamori, H., ed. (2009). *Earthquake Seismology* (Treatise on Geophysics vol. 4). Elsevier, Amsterdam. 700 pp.
- Kanamori, H. and E. Boschi, eds. (1983). *Earthquakes: Observation, Theory and Interpretation*. North-Holland, Amsterdam. 608 pp.

- Karnik, V. (1969, 1971). *Seismicity of the European Area* (2 vols.). D. Reidel, Dordrecht. 215 pp. and 223 pp.
- Kasahara, K. (1981). *Earthquake Mechanics*. Cambridge University Press, Cambridge. 433 pp.
- Kennett, B. L. N. (1983). *Seismic Wave Propagation in Stratified Media*. Cambridge University Press, Cambridge. 342 pp.
- Kennett, B. L. N. (2001). *The Seismic Wave Field. Vol 1: Introduction and Theoretical Development*. Cambridge University Press, Cambridge. 371 pp.
- Kennett, B. L. N. (2003). *The Seismic Wave Field. Vol. 2: Interpretation of Seismograms on Regional and Global Scales*. Cambridge University Press, Cambridge. 541 pp.
- Kisslinger, C. and Z. Zuzuki, eds. (1977). *Earthquakes Precursors*. Center Academic Publisher, Tokyo.
- Kostrov, B. V. and S. Das (1988). *Principles of Earthquake Source Mechanics*. Cambridge University Press, Cambridge. 286 pp.
- Koyama, J. (1997). *The Complex Faulting Process of Earthquakes*. Kluwer Academic, Dordrecht. 194 pp.
- Kulhanek, O. (1990). *Anatomy of Seismograms*. Elsevier, Amsterdam. 178 pp.
- Lapwood, E. R. and T. Usami (1981). *Free Oscillations of the Earth*. Cambridge University Press, Cambridge. 243 pp.
- Lomnitz, C. (1974). *Global Tectonics and Earthquake Risk*. Elsevier, Amsterdam. 319 pp.
- Lomnitz, C. (1994). *Fundamentals of Earthquake Prediction*. John Wiley, New York. 326 pp.
- Love, A. E. H. (1911). *Some Problems of Geodynamics*. Cambridge University Press, Cambridge. 181 pp.
- Meissner, R. (1986). *The Continental Crust, A Geophysical Approach*. Academic Press, San Diego.
- Nolet, G., ed. (1987). *Seismic Tomography with Applications in Global Seismology and Exploration Geophysics*. D. Reidel, Dordrecht. 386 pp.
- Officer, C. B. (1958). *Introduction to the Theory of Sound Transmission*. McGraw-Hill, New York. 284 pp.
- Payo, G. (1986). *Introducción al análisis de sismogramas*. Instituto Geográfico Nacional, Madrid. 175 pp.
- Pilant, W. L. (1979). *Elastic Waves in the Earth*. Elsevier, Amsterdam. 493 pp.
- Reiter, C. (1990). *Earthquake Hazard Analysis, Issues and Insights*. Columbia University Press, New York. 254 pp.
- Rikitake, T. (1982). *Earthquake Forecasting and Warning*. D. Reidel, Dordrecht. 402 pp.
- Romanowicz and A. Dziewonski, eds. (2015). *Seismology and Structure of the Earth*. Elsevier, Amsterdam. 858 pp.
- Sato, H. and M. C. Fehler (1998). *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*. Springer, New York. 308 pp.
- Sawarsky, E. (1975). *Seismic Waves*. Mir, Moscow. 350 pp.
- Scherbaum, F. (1996). *Of Poles and Zeros. Fundamentals of Digital Seismology*. Kluwer Academic, Dordrecht. 256 pp.
- Scholz, C. H. (1990, 2002). *The Mechanics of Earthquakes and Faulting*. Cambridge University Press, Cambridge. 433 pp.
- Sheriff, R. E. and L. P. Geldart (1982, 1983, 1995). *Exploration Seismology* (2 volumes). Cambridge University Press, Cambridge. 1st 272 pp. and 240 pp. 2nd 628 pp.

- Simpson, D. W. and P. G. Richards, eds. (1981). *Earthquake Prediction: An International Review*. American Geophysical Union, Washington.
- Tong, V. C. H. and R. A. Garcia (2015). *Extraterrestrial Seismology*. Cambridge University Press, Cambridge. 441 pp.
- Udias, A., R. Madariaga, and E. Buforn (2014). *Source Mechanism of Earthquakes. Theory and Practice*. Cambridge University Press, Cambridge. 302 pp.
- Vvedenskaia, A. V. (1984). *Seismodinamika*. Isdatelsvo "Nauka", Moscow. 144 pp.

Elasticity and wave mechanics

- Achenbach, J. D. (1973). *Wave Propagation in Elastic Solids*. North-Holland, Amsterdam. 425 pp.
- Bedford, A. and D. S. Drumheller (1996). *Introduction to Elastic Wave Propagation*. Wiley, New York. 312 pp.
- Brekhovskikh, L. and V. Goncharov (1985). *Mechanics of Continua and Wave Dynamics*. Springer, Berlin. 342 pp.
- Eringen, A. C. and E. S. Suhubi (1974–1975) *Elastodynamics* (Vol. I and II) Academic Press, New York.
- Kolymbas, D. and G. Viggiani, eds. (2009) *Mechanics of Natural Solids*. Springer, Berlin. 297 pp.
- Lanczos, C. (1986). *The Variational Principles of Mechanics*. Dover Publications, New York. 418 pp.
- Landau, L. D. and E. M. Lifshitz (1970). *Theory of Elasticity*. Pergamon Press, Oxford. 134 pp.
- Leipholz, H. (1974). *Theory of Elasticity*. Noordhoff, Leiden. 400 pp.
- Lindsay, R. B. (1960). *Mechanical Radiation*. McGraw-Hill, New York. 415 pp.
- Love, A. E. H. (1892, 1893, 1906, 1927). *The Mathematical Theory of Elasticity*. Cambridge University Press, Cambridge. 641 pp.
- Malvern, L. E. (1969). *Introduction to the Mechanics of a Continuous Medium*. Prentice Hall, Englewood Cliffs. 713 pp.
- Sommerfeld, A. (1964). *Mechanics of Deformable Bodies*. Academic Press, New York. 396 pp.
- Timoshenko, I. S. and J. N. Goodier (1951). *Theory of Elasticity*. McGraw-Hill, New York. 506 pp.
- Tolstoy, I. (1973). *Wave Propagation*. McGraw-Hill, New York. 458 pp.

References

Please note that citations in the text may refer to books listed in the Bibliography. These are not repeated here.

- Aki, K. (1966). Generation and propagation of G waves from Niigata earthquake of June 16, 1964. Estimation of earthquake movement, released energy and stress-strain drop from G wave spectrum. *Bull. Earthq. Res. Inst.* **44**, 23–88.
- Aki, K. (1967). Scaling law of seismic spectrum. *J. Geophys. Res.* **73**, 5359–76.
- Aki, K. (1969). Analysis of seismic coda of local earthquakes as scattered waves. *J. Geophys. Res.* **74**, 615–31.
- Aki, K. (1979). Characterization of barriers on an earthquake fault. *J. Geophys. Res.* **84**, 6140–8.
- Aki, K. (1981). A probability synthesis of precursory phenomena. In D. W. Simpson and P. G. Richards (eds.) *Earthquake Prediction: An International Review*. American Geophysical Union, Washington. pp. 566–74.
- Aki, K. (1984). Asperities, barriers, characteristic earthquakes and strong motion prediction. *J. Geophys. Res.* **89**, 5867–72.
- Aki, K. (1988). Physical theory of earthquakes. In J. Bonnin, M. Cara, A. Cisternas, and R. Fantechi (eds.) *Seismic Hazard in Mediterranean Region*. Kluwer Academic, Dordrecht, pp. 566–74.
- Aki, K. and B. Chouet (1975). Origin of coda waves: source, attenuation and scattering effects. *J. Geophys. Res.* **80**, 3322–42.
- Aki, K., A. Christoffersson, and E. S. Husebye (1976). Determination of the three-dimensional seismic structures of the lithosphere. *J. Geophys. Res.* **82**, 277–96.
- Allen, R.M., P. Gasparini, O. Kamigaichi, and M. Bose (2009). The status of earthquake early warning around the world: an introductory overview. *Seism. Res. Lett.* **80**, 682–93 doi: 10.1785/gssrl.80.5.682
- Alinghi, A. (2003). *Receiver function analysis of the crust and upper mantle from the North German Basin to Tharheim Baltic Shield*. Doctoral dissertation, Freie Universität, Berlin.
- Alsop, L. E., G. H. Sutton, and M. Ewing (1961). Free oscillations of the Earth observed on strain and pendulum seismographs. *J. Geophys. Res.* **66**, 631–41.
- Anderson, D. L. and C. B. Archambeau (1964). The anelasticity of the Earth. *J. Geophys. Res.* **69**, 2071–84.
- Anderson, D. L. and R. S. Hart (1978). Q of the Earth. *J. Geophys. Res.* **83**, 5869–82.
- Andrews, D. J. (1980). A stochastic fault model. I, static case. *J. Geophys. Res.* **85**, 3867–77.
- Andrews, D. J. (1989). Mechanics of fault junctions. *J. Geophys. Res.* **94**, 9389–97.

- Archuleta, R. J. (1984). A faulting model for the 1970 Imperial Valley earthquake. *J. Geophys. Res.* **89**, 4559–85.
- Arias, A. (1970). A measure of earthquake intensity. In R. J. Hansen (ed.) *Seismic Design for Nuclear Power Plants*. MIT Press, Cambridge (Mass.), 438–83.
- Atkinson, G. M. and S. I. Kaka (2007). Relationships between felt intensity and instrumental ground motion in the central United States and California. *Bull. Seismol. Soc. Am.* **97**, 497–510.
- Backus, G. and M. Mulcahy (1976a). Moment tensors and other phenomenological descriptions of seismic sources. I Continuous displacements. *Geophys. J. R. Astr. Soc.* **46**, 341–61.
- Backus, G. and M. Mulcahy (1976b). Moment tensors and other phenomenological descriptions of seismic sources. II Discontinuous displacements. *Geophys. J. R. Astr. Soc.* **47**, 301–30.
- Bak, P. and C. Tang (1989). Earthquakes as a self-organized critical phenomenon. *J. Geophys. Res.* **94**, 15635–7.
- Bakun, W. H. and C. M. Wentworth (1997). Estimating earthquake location and magnitude from seismic intensity data. *Bull. Seismol. Soc. Am.* **87**, 1502–21.
- Balakina, L. M., E. F. Savarensky, and A. V. Vvedenskaya (1961). On determination of earthquake mechanism. *Phys. Chem. Earth* **4**, 211–38.
- Barenblatt, G. I. (1959). The formation of equilibrium cracks during brittle fracture. General ideas and hypothesis. *J. Appl. Math. Mech.* **23**, 622–36.
- Barriere, B. and D. L. Turcotte (1991). A scale-invariant cellular automata model for distributed seismicity. *Geophys. Res. Lett.* **18**, 2011–14.
- Benioff, H., F. Press, and S. Smith (1961). Excitation of the free oscillations of the Earth by earthquakes. *J. Geophys. Res.* **66**, 605–19.
- Ben Menahem, A. (1961). Radiation of seismic surface waves from finite moving sources. *Bull. Seismol. Soc. Am.* **51**, 401–53.
- Ben Menahem, A. (1962). Radiation of seismic body waves from finite moving sources in the earth. *J. Geophys. Res.* **67**, 396–474.
- Berckhemer, H. (1962). Die Ausdehnung der Bruchfläche im Erdbebenherd und ihr Einfluß auf das seismische Wellenspektrum. *Gerland Beitr. Geophys.* **71**, 5–26.
- Bisztricsány, E. A. (1958). A new method for determination of magnitude of earthquakes (in Hungarian). *Geofizikai Közlemények* **6**, 69–76.
- Blanco, M. J. and W. Spakman (1993). The P wave velocity structure of the mantle below the Iberian peninsula: evidence for a subducted lithosphere below southern Spain. *Tectonophysics* **221**, 13–34.
- Bolt, B. A. (1960). Earthquake epicenter, focal depth and origin time using a high speed computer. *Geophys. J. R. Astr. Soc.* **3**, 433–40.
- Bolt, B. A. (1982). The constitution of the core: seismological evidence. *Phil. Trans. R. Soc. A* **306**, 11–20.
- Bolt, B. A. and R. A. Uhrhammer (1981). The structure, density and homogeneity of the Earth's core. In *Evolution of the Earth*, Geodynamics Series, American Geophysical Union, Washington. Volume **5**, pp. 28–37.

- Bommer, J. J. (2002). Deterministic vs probabilistic seismic hazard assessment: an exaggerated and obstructive dichotomy. *Jour. Earthquake Engineering* **6**, 43–73.
- Bonito, M. (1691). *Terra tremante o vero continuatione de terremoto*, A. Parrino e M. L. Mutti, Naples, Italy.
- Boore, D. M., C. Stephens, and W. B. Joyner (2002). Comments on baseline correction of digital strong-motion data: examples from the 1999 Hector Mine, California, earthquake. *Bull. Seismol. Soc. Am.*, **92**, 1543–60.
- Bormann, P. and J. W. Dewey (2014). The new IASPEI standards for determining magnitudes from digital data and their relation to classical magnitudes. In P. Bormann (ed.), *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, Deutsches GeoForschungsZentrum, Potsdam, 1–44. DOI: http://doi.org/10.2312/GFZ.NMSOP-2_IS_3.3
- Boyd, O. S. and C. H. Cramer (2014). Estimating earthquake magnitudes from reported intensities in the Central and Eastern United States. *Bull. Seismol. Soc. Am.*, **104**, 1709–22.
- Bracewell, R. N. (1986). *The Fourier Transform and its Applications* (2nd revised edn). McGraw-Hill International, New York. 474 pp.
- Brillinger, D., A. Udías, and B. A. Bolt (1980). A probability model for regional focal mechanism solutions. *Bull. Seismol. Soc. Am.* **70**, 149–70.
- Brune, J. N. (1970). Tectonic stress and the spectra of shear waves from earthquakes. *J. Geophys. Res.* **75**, 4997–5009.
- Brune, J. N., J. E. Nafe, and L. Alsop (1961). The polar phase shift of surface waves on a sphere. *Bull. Seismol. Soc. Am.* **51**, 247–57.
- Bufo, E. (1994). Métodos para la determinación del mecanismo focal de los terremotos. *Física de la tierra* **6**, 113–40.
- Buland, R., J. Berger, and F. Gilbert (1979). Observations from IDA network of attenuation and splitting during a recent earthquake. *Nature* **277**, 358–62.
- Burridge, R. (1969). The numerical solution of certain integral equations with non-integrable kernels arising in the theory of cracks propagation and elastic wave diffraction. *Phil. Trans. R. Soc. A* **265**, 353–81.
- Burridge, R. and L. Knopoff (1964). Body force equivalents for seismic dislocations. *Bull. Seismol. Soc. Am.* **54**, 1875–88.
- Burridge, R. and L. Knopoff (1967). Model and theoretical seismicity. *Bull. Seismol. Soc. Am.* **57**, 341–71.
- Byerly, P. (1928). The nature of the first motion in the Chilean earthquake of November 11, 1922. *Am. J. Sci.* **16**, 232–6.
- Cesca, S., S. Heimann, K. Stammler, and T. Dahm (2010). Automated procedure for point and kinematic source inversion at regional distances. *J. Geophys. Res.* **115**, B06304, DOI: 10.1029/2009JB006450, 2010
- Cisternas, A. (1982). *Notas del Curso de Sismología*. la. Escuela Internacional de Riesgo Sísmico, ACIF, Bogotá.
- Cochard, A. and R. Madariaga (1994). Dynamic faulting under rate dependent friction. *Pageophysics* **142**, 419–45.
- Coen, D. R. (2013). *The Earthquake Observers: Disaster Science from Lisbon to Richter*. The University of Chicago Press, Chicago. 348 pp.

- Cornell, C. A. (1968). Engineering risk analysis. *Bull. Seismol. Soc. Am.* **58**, 1583–1606.
- Crampin, S. (1977). A review of the effects of anisotropic layering on the propagation of seismic waves. *Geophys. J. R. Astr. Soc.* **49**, 9–27.
- Crampin, S. (1978). Seismic wave propagation through a cracked solid: Polarization as a possible dilatancy diagnostic. *Geophys. J. R. Astr. Soc.* **53**, 467–96.
- Dahlen, F. A. (1980). Splitting of the oscillations of the Earth. In A. M. Dziewonski and E. Boschi (eds.) *Physics of the Earth Interior*, North-Holland, Amsterdam, pp. 82–126.
- Dainty, A. M. (1981). A scattering model to explain seismic Q observations in the lithosphere between 1 and 30 Hz. *Geophys. Res. Lett.* **8**, 1126–8.
- Das, S. and K. Aki (1977). Fault planes with barriers: A versatile earthquake model. *J. Geophys. Res.* **82**, 5648–70.
- Derr, J. S. (1969). Free oscillations observations through 1968. *Bull. Seismol. Soc. Am.* **59**, 2079–99.
- Deschamps, A., H. Lyon-Caen, and R. Madariaga (1980). Mise au point sur les méthodes de calcul de séismogrammes synthétiques de long période. *Ann. Géophys.* **36**, 167–78.
- Dewey, J. W. (1972). Seismicity and tectonics of western Venezuela. *Bull. Seismol. Soc. Am.* **62**, 1711–51.
- Dewey, J. W. and P. Byerly (1969). The early history of seismometry (to 1900). *Bull. Seismol. Soc. Am.* **59**, 183–227.
- Dieterich, J. H. (1972). Time dependence friction as a possible mechanism for aftershocks. *J. Geophys. Res.* **77**, 3771–81.
- Douglas, A. (1967). Joint epicenter determination. *Nature* **215**, 47–8.
- Dziewonski, A. M. and D. L. Anderson (1981). Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356.
- Dziewonski, A. M. and F. Gilbert (1972). Observation of normal modes from 84 recordings of the Alaskan earthquake of 1964, March 28. *Geophys. J. R. Astr. Soc.* **27**, 393–446.
- Dziewonski, A. M. and A. L. Hales (1972). Numerical analysis of dispersed seismic waves. In B. A. Bolt (ed.) *Methods in Computational Physics, Volume 11, Seismology: Surface Waves and Earth Oscillations*. Academic Press, New York. pp. 39–85.
- Dziewonski, A. M. and J. H. Woodhouse (1987). Global images of the Earth's interior. *Science* **236**, 37–48.
- Dziewonski, A. M., T. A. Chou, and J. H. Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res.* **86**, 2825–52.
- Espinosa, A. F. and A. Lopez-Arroyo (1977). Earthquake instrumental intensity from strong ground motion records: San Fernando earthquake. *Proceedings of the 6th World Conference on Earthquake Engineering, New Delhi, India* (Vol. **2**, pp. 1–7).
- Evans, R. (1997). Assessment of schemes for earthquake prediction: Editor introduction. *Geophys. J. Int.* **131**, 413–20.
- Ferrari, G. (ed.) (1992). *Two Hundred Years of Seismic Instruments in Italy, 1731–1940*. Istituto Nazionale di Geofisica, Rome. 156 pp.
- Freund, L. B. (1972). Crack propagation in an elastic solid subjected to general loading. I. Constant rate of extension. *J. Mech. Phys. Solids* **20**, 129–40.

- Freund, L. B. (1979). The mechanics of dynamic shear crack propagation. *J. Geophys. Res.* **84**, 2199–209.
- Fukuyama, E. and K. Irikura (1986). Rupture process of the 1983 Japan Sea (Akita-Oki) earthquake using a waveform inversion method. *Bull. Seismol. Soc. Am.* **76**, 1623–49.
- Gallart, J., N. Vidal, A. Estévez, *et al.* (1995). The ESCI-Valencia Trough vertical reflection experiment: A seismic image of the crust from NE Iberian Peninsula to the western Mediterranean. *Rev. Soc. Geol. España* **8**, 401–15.
- Geiger, L. (1910). Herdbestimmung bei Erdbeben aus Ankunftszeiten. *Nachrichten Kön. Gesell. der Wissenschaft Göttingen, Math. Phys. K* 331–49.
- Geller, R. J. (1997). Earthquake prediction. A critical review. *Geophys. J. Int.* **131**, 425–50.
- Georgudas, I.G., G. Ch. Sirakoulis, and I. Andreadis (2007). Modelling earthquake activity features using cellular automata. *Math. and Comp. Modelling* **46**, 124–37.
- Gilbert, F. (1970). Excitation of the normal modes of the Earth by earthquakes sources. *Geophys. J. R. Astr. Soc.* **22**, 223–6.
- Gilbert, F. and G. E. Backus (1966). Propagator matrices in elastic wave and vibration problems. *Geophysics* **31**, 326–32.
- Gilbert, F. and A. M. Dziewonski (1975). An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra. *Phil. Trans. R. Soc. A* **214**, 369–71.
- Grant, S. P. (2002). Mantle shear-wave tomography and the fate of subducted slabs. *Phil. Trans. Roy. Soc. London, A* **360**, 2475–91.
- Grünthal, G. (ed.) (1993). European Macroseismic Scale 1992 (up-dated MSK scale). *Cahiers du Centre Européen de Géodynamique et de Seismologie*, Luxembourg. Volume **7**, 79 pp.
- Grünthal, G. (ed.) (1998). European Macroseismic Scale 1998. *Cahiers du Centre Européen de Géodynamique et de Seismologie*, Volume **15**. Luxembourg. 99 pp.
- Gubbins, D., B. Sreenivasan, J. Mound, and S. Rost (2011). Melting of the Earth's inner core. *Nature* **473**(7347), 361–3.
- Gung, Y., M. Panning, and B. Romanowicz (2003). Global anisotropy and the thickness of continents. *Nature* **442**, 707–11.
- Gutdeutsch, R., D. Kaiser, and G. Jentzsch (2002). Estimation of earthquake magnitude from epicentral intensities and other focal parameters in central and southern Europe. *Geophys. J. Int.* **151**, 824–834.
- Gutenberg, B. (1945). Amplitudes of surface waves and magnitudes of shallow earthquakes. *Bull. Seismol. Soc. Am.* **35**, 3–12.
- Gutenberg, B. and C. F. Richter (1942). Earthquake magnitude, intensity, energy and acceleration. *Bull. Seismol. Soc. Am.* **32**, 163–91.
- Gutenberg, B. and C. F. Richter (1956). Magnitude and energy of earthquakes. *Annali di geofisica* **9**, 1–15.
- Hanks, T. C. (1982). f_{max} . *Bull. Seismol. Soc. Am.* **72**, 1867–79.
- Hanks, T. C. and H. Kanamori (1979). A moment magnitude scale. *J. Geophys. Res.* **84**, 2348–50.
- Hanks, T. C. and M. Wyss (1972). The use of body wave spectra in the determination of seismic source parameters. *Bull. Seismol. Soc. Am.* **62**, 561–89.

- Hardebeck and Shearer (2002). A new method for determining first-motion focal mechanisms. *Bull. Seismol. Soc. Am.* **92**, 2264–76.
- Hartzell, S. (1978). Earthquake aftershocks as Green's functions. *Geophys. Res. Lett.* **5**, 1–4.
- Hartzell, S. H. (1989). Comparison of seismic waveforms inversion results for the rupture history of a finite fault. Application to the 1986 North Palm Springs, California earthquake. *J. Geophys. Res.* **94**, 7515–34.
- Hartzell, S. H. and T. Heaton (1983). Inversion of strong ground motion and teleseismic wave form data for the fault rupture history of the 1970 Imperial Valley, California earthquake. *Bull. Seismol. Soc. Am.* **73**, 1553–83.
- Haskell, N. A. (1953). The dispersion of surface waves on multilayered media. *Bull. Seismol. Soc. Am.* **43**, 17–34.
- Haskell, N. A. (1964). Total energy and energy spectral density of elastic wave radiation from propagating faults. *Bull. Seismol. Soc. Am.* **54**, 1811–41.
- Haskell, N. A. (1966). Total energy and energy spectral density of elastic wave radiation from propagating faults. Part II. *Bull. Seismol. Soc. Am.* **56**, 125–40.
- Havskov, J. and L. Ottemöller (2010). *Routine Data Processing in Earthquake Seismology*. Springer, Berlin.
- Hayakawa, M. and Y. Hobara (2010). Current status of seismic electromagnetics for short-term earthquake prediction. *Geomatics Nat. Hazards Risk* **1**, 115–55.
- Heaton, T. H. (1990). Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Phys. Earth Planet. Inter.* **64**, 1–20.
- Helffrich, G., J. Wookey, and I. Bastow (2013). *The Seismic Analysis Code. A Primer and User's Guide*. Cambridge University Press, Cambridge.
- Herrmann, R. B. (2013) Computer programs in seismology: An evolving tool for instruction and research, *Seism. Res. Lettr.* **84**, 1081–8, doi:10.1785/0220110096
- Herráiz, M. and A. F. Espinosa (1987). Coda waves: A review. *Pageoph.* **125**, 499–577.
- Herrin, E. (1968). 1968 Seismological tables for P phases. *Bull. Seismol. Soc. Am.* **58**, 1193–241.
- Hirata, T. (1989). A correlation between the b value and the fractal dimension of earthquakes. *J. Geophys. Res.* **94**, 7507–14.
- Hough, S. E. (2010). *Predicting the Unpredictable: The Tumultuous Science of Earthquake Prediction*. Princeton University Press, Princeton (NJ), 280 pp.
- Ida, Y. (1972). Cohesive force across the tip of a longitudinal shear crack and Griffith's specific surface energy. *J. Geophys. Res.* **77**, 3796–805.
- Igel, H., N. Takeuchi, R. Geller, *et al.* (2000). The COSY Project: verification of global seismic modeling algorithms. *Phys. Earth Plan. Int.* **119**, 3–23.
- Inglada, V. (1928). Die Berechnung der Herdkoordinaten eines Nahbebens aus den Eintrittszeiten der in einigen benachbarten Stationen auf gezeichneten P oder S Wellen. *Beiträge zur Geophysik* **19**, 73–98.
- Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata (1990). Whole mantle P wave travel time tomography. *Phys. Earth Planet. Inter.* **59**, 294–328.
- Isacks, B. L., J. Oliver, and L. R. Sykes (1968). Seismology and new global tectonics. *J. Geophys. Res.* **73**, 5855–99.

- Ito, K. and M. Matsuzaki (1990). Earthquakes as self-organized critical phenomena. *J. Geophys. Res.* **95**, 6853–60.
- Jackson, D. D. and D. L. Anderson (1970). Physical mechanism of seismic attenuation. *Rev. Geophys. Space Phys.* **8**, 1–63.
- Jeanloz, R. (1990). The nature of the Earth's core. *Ann. Rev. Earth Planet. Sci.* **18**, 357–86.
- Jeffreys, H. (1931). On the cause of oscillatory movements in seismograms. *Mon. Not. R. Astr. Soc. Geophys. Suppl.* **2**, 407–16.
- Jeffreys, H. (1957). The damping of S waves. *Nature* **207**, 675.
- Jeffreys, H. and K. H. Bullen (1940). *Seismological Tables*. British Association for the Advancement of Science, London. 50 pp.
- Jones, L. (2015). Resilience by Design: Bringing Science to Policy Makers. *Seism. Res. Lett.* **86**, 294–301, doi: 10.1785/0220150010
- Jost, M. L. and R. B. Herrmann (1989). A student's guide to and review of moment tensors. *Seism. Res. Lett.* **60**, 37–57.
- Kanamori, H. (1977). The energy release in great earthquakes. *J. Geophys. Res.* **82**, 2921–87.
- Kanamori, H. (2003). Earthquake prediction: An overview. *International Handbook of Earthquake and Engineering Seismology*. Academic Press, San Diego CA, 1205–16.
- Kanamori, H. and D. L. Anderson (1975). Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* **65**, 1073–95.
- Kanamori, H. and J. W. Given (1981). Use of long period surface waves for rapid determination of earthquake source parameters. *Phys. Earth Planet. Inter.* **27**, 8–31.
- Kanamori, H. and G. S. Stewart (1978). Seismological aspects of the Guatemala earthquake of February 4, 1976. *J. Geophys. Res.* **83**, 3427–34.
- Kasahara, K. (1963). Computer program for fault-plane solutions. *Bull. Seismol. Soc. Am.* **53**, 1–13.
- Kearey, P. and F. J. Vine (1996). *Global Tectonics* (2 ed.). Blackwell Science, Oxford.
- Kendall, J. M. (2000) Seismic anisotropy in the boundary layers of the mantle, in Earth's deep interior. In S. I. Karato, A. Forte, R. Liebermann, G. Masters, and L. Stixrude (eds.) *Mineral Physics and Tomography from the Atomic to the Global Scale*. American Geophysical Union, Washington, D. C. doi: 10.1029/GM117p0133
- Kennett, B. L. N. and E. R. Engdahl (1991). Travel times for global earthquake location and phase identification. *Geophys. J. Int.* **105**, 429–65.
- Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from travel times. *Geophys. J. Int.* **122**, 108–24.
- Keylis-Borok, V. I. (1956). Methods and results of the investigation of earthquake mechanism. *Publ. Bureau Central Sismol. Int. Ser. A. Trav. Scient.* **19**, 383–94.
- Keylis-Borok, V. I. (1959). On the estimation of the displacement in an earthquake source and of source dimensions. *Ann. Geophys.* **12**, 205–14.
- Kikuchi, A. and H. Kanamori (1982). Inversion of complex body waves. *Bull. Seismol. Soc. Am.* **72**, 491–506.
- Kikuchi, A. and H. Kanamori (1991). Inversion of complex waves III, *Bull. Seismol. Soc. Am.* **81**, 2335–50.

- Kind, R., X. Yuan, J. Saul, *et al.* (2002). Seismic images of the crust and upper mantle beneath Tibet: evidence for Eurasian Plate subduction. *Science* **298**, 1219–21.
- Klein, F. (2000). Hypocenter location program HYPOINVERSE. Users Guide to version 2000. U. S. Geol. Survey. Open File Report, version 0.99.
- Knopoff, L. (1961). Analytical calculation of the fault plane problem. *Publ. Dom. Obs. (Ottawa)* **24**, 309–15.
- Knopoff, L. (1964). A matrix method for elastic wave problems. *Bull. Seismol. Soc. Am.* **54**, 431–8.
- Knopoff, L. (1967). *The Mathematics of the Seismic Source*. Institut für Geophysik, Universität Karlsruhe (lecture notes).
- Knopoff, L. (1972). Observation and inversion of surface wave dispersion. *Tectonophysics* **13**, 497–519.
- Knopoff, L. and F. Gilbert (1960). First motions from seismic sources. *Bull. Seismol. Soc. Am.* **50**, 117–34.
- Knopoff, L. and M. Randall (1970). The compensated vector linear dipole: a possible mechanism for deep earthquakes. *J. Geophys. Res.* **75**, 4957–63.
- Kostrov, B. V. (1964). Self-similar problems of propagation of shear cracks. *J. Appl. Math. Mech.* **28**, 1077–87.
- Kostrov, B. V. (1966). Unsteady propagation of longitudinal cracks. *J. Appl. Math. Mech.* **30**, 1241–8.
- Lamb, H. (1904). On the propagation of tremors over the surface of an elastic solid. *Phil. Trans. Roy. Soc. A* **203**, 1–42.
- Lanczos, C. (1957). *Applied Analysis*. Isaac Pitman, London. 539 pp.
- Landisman, M., A. Dziewonski, and Y. Satô (1969). Recent improvements in the analysis of surface waves observations. *Geophys. J. R. Astr. Soc.* **17**, 369–403.
- Langston, C. A. and D. V. Helmberger (1975). A procedure for modelling shallow dislocation sources. *Geophys. J. R. Astr. Soc.* **42**, 117–30.
- Langston, C. A. (1977). The effect of planar dipping structure on source and receiver for constant ray parameter responses. *Bull. Seismol. Soc. Am.* **67**, 1029–50.
- Larson, K. (2009). GPS seismology. *J. Geod.* **83**, 227–33.
- Lawrence, J. F. and M. E. Wyss (2006) QLM9: A new radial quality factor (Q_μ) model for the lower mantle. *Earth and Plan. Sci. Lett.* **241**, 962–71.
- Lee, W. B. and S. C. Solomon (1978). Simultaneous inversion of surface wave phase velocity and attenuation: Love waves in western North America. *J. Geophys. Res.* **83**, 3389–400.
- Lee, W. H. K. and J. C. Lahr (1971). HYPO71: A computer program for determining hypocenter, magnitude and first motion pattern of local earthquakes. U.S. Geological Survey, Menlo Park. Open-file report, pp. 75–311.
- Lee, W. H. K., R. B. Bennet, and K. L. Meahger (1972). A method of estimating magnitude of local earthquakes from signal duration. U.S. Geological Survey, Menlo Park. Open-file report.
- Leveque, J. J. and M. Cara (1985). Inversion of multimode surface wave data: Evidence for sublithospheric anisotropy. *Geophys. J. R. Astr. Soc.* **83**, 753–74.

- Lindh A., P. Evans, P. Harsh, and G. Buhr (1979). The Parkfield experiment. *Earthq. Inf. Bull.* **11**, 205–13.
- Lomax, A. (2005). A reanalysis of the hypocentral location and related observations for the Great 1906 California Earthquake. *Bull. Seismol. Soc. Am.*, **91**, 861–77.
- Lomnitz, C. (1957). Linear dissipation in solids. *J. Appl. Phys.* **28**, 201–5.
- Macelwane, J. B. (1946). Forecasting earthquakes. *Bull. Seismol. Soc. Am.* **36**, 1–4.
- MacGaffrey, R., G. Abers, and P. Zwick (1991). Inversion of teleseismic body waves. In W. H. K. Lee (ed.). *Digital Seismogram Analysis and Wave Form Inversion. IASPEI Software Library*. Volume **3**, pp. 81–166.
- Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bull. Seismol. Soc. Am.* **66**, 639–66.
- Madariaga, R. (1979). On the relation between seismic moment and stress drop in the presence of stress and strength heterogeneity. *J. Geophys. Res.* **84**, 2243–50.
- Madariaga, R. (1983). Earthquake source theory. A review. In H. Kanamori and E. Boschi (eds.) *Earthquakes Observations, Theory and Interpretation*. North-Holland, Amsterdam, pp. 1–44.
- Madariaga, R. (1989). Propagación de ondas sísmicas en el campo cercano. *Física de la tierra* **1**, 51–73.
- Mallet, R. (1848). On the dynamics of earthquakes; being an attempt to reduce their observed phenomena to the known laws of wave motion in solids and fluid. *Trans. Royal Irish Academy* **21**, 51–105.
- Mallet, R. (1862) *Great Neapolitan Earthquake of 1857. The First Principles of Observational Seismology*. Chapman and Hall, London.
- Mallet, R. and J. W. Mallet (1858). *Earthquake Catalogue of the British Association with the Discussion, Curves and Maps*. Taylor and Francis, London, United Kingdom.
- Mandelbrot, B. (1977). *Fractals: Form, Chance and Dimension*. W. H. Freeman, San Francisco.
- Marquardt, D. W. (1963). An algorithm for least squares estimation of non linear parameters. *J. Soc. Ind. Appl. Math.* **11**, 431–41.
- Mattesini, M., A. B. Belonoshko, E. Buforn, *et al.* (2010). Hemispherical anisotropic patterns of the Earth's inner core. *Proc. Natl. Acad. Sci.* **107**(21), 9507–12.
- Mattesini, M., A. B. Belonoshko, H. Tkalčić, *et al.* (2013). Candy wrapper for the Earth's inner core. *Sci. Rep.* **3**, 2096, doi:10.1038/srep02096.
- Maupin, V. and J. Park. (2009). Theory and observations. Wave propagation in anisotropic media. In B. Romanowicz and A. Dziewonski (eds.). *Seismology and Structure of the Earth*. Elsevier, Amsterdam, 289–322.
- McKenzie, D. P. (1972). Active tectonics of the Mediterranean region. *Geophys. J. R. Astr. Soc.* **30**, 109–85.
- Melton, B. S. (1981). Earthquake seismograph development: A modern history. *EOS (Trans. Am. Geophys. U.)* **62**, 505–10, 545–8.
- Mendiguren, L. A. (1977). Inversion of surface wave data in source mechanism studies. *J. Geophys. Res.* **82**, 889–94.

- Mendoza, C. and S. Hartzell (1989). Slip distribution of the 19 September 1985 Michoacan, Mexico, earthquake: Near source and teleseismic constraints. *Bull. Seismol. Soc. Am.* **79**, 655–69.
- Merz, H. and C. A. Cornell (1973). Seismic risk analysis based on a quadratic magnitude frequency law. *Bull. Seismol. Soc. Am.* **63**, 1999–2006.
- Mikumo, T. and T. Miyatake (1979). Earthquake sequences on a frictional fault model with non-uniform strengths and relaxation times. *Geophys. J. R. Astr. Soc.* **54**, 417–38.
- Milne, J. (1911). *Catalogue of Destructive Earthquakes*. Burlington House, London, United Kingdom.
- Minster, J. B. (1980). Anelasticity and attenuation. In A. M. Dziewonski and E. Boschi (eds.), *Physics of the Earth Interior*. North-Holland, Amsterdam, pp. 152–212.
- Misra, P. and P. Enge (2012). *Global Positioning System. Signal, Measurements, and Performance*. (Revised Second Edition). Ganga-Jamuna Press, Lincoln (Mass).
- Mogi, K. (1963). Some discussions on aftershocks, foreshocks and earthquake swarms, the fracture of a medi-infinite body caused by inner stress origin and its relation to earthquake phenomena. *Bull. Earthq. Res. Inst. Tokyo Univ.* **41**, 615–58.
- Molnar, P., B. E. Tucker, and J. N. Brune (1973). Corner frequencies of P and S waves and models of earthquake sources. *Bull. Seismol. Soc. Am.* **63**, 2091–104.
- Moreira de Mendonça, J. J. (1758). *Historia universal dos terremotos*, Antonio Vicenta da Silva, Lisbon, Portugal.
- Morelli, A. and A. M. Dziewonski (1993). Body wave travel times and spherically symmetric P and S wave velocity model. *Geophys. J. Int.* **112**, 178–94.
- Morelli, A., A. M. Dziewonski, and J. H. Woodhouse (1986). Anisotropy of the inner core inferred from PKiKP travel times, *Geophys. Res. Lett.* **13**(13), 1545–8.
- Morse, P. M. and H. Feshbach (1953). *Methods of Theoretical Physics*. McGraw-Hill, New York.
- Mualchin, L. (2010). History of modern earthquake hazard mapping and assessment in California using deterministic or scenario approach. *Pure App. Geophys.* **168**, 383–407.
- Müller, G. (1985). The reflectivity method: a tutorial. *J. Geophys.* **58**, 153–74.
- Nabelek, J. L. (1984). *Determination of earthquake source parameters from inversion of body waves*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Nakano, H. (1923). Notes on the nature of forces which give rise to the earthquake motion. *Seis. Bull. Cent. Meteor. Obs. Japan* **1**, 92–120.
- Nishimura, C. E. and D. W. Forsyth (1989). The anisotropic structure of the upper mantle in the Pacific. *Geophys. J. Int.* **96**, 203–29.
- Nishimura, G. (1937). On the elastic waves due to pressure variations on the inner surface of a spherical cavity in an elastic solid. *Bull. Earthq. Res. Inst. Tokyo* **15**, 614–35.
- Nuttli, O. W. (1974). Seismic wave attenuation and magnitude relations for east and north America. *J. Geophys. Res.* **78**, 876–85.
- Okal, E. (1992). A student's guide to teleseismic body wave amplitudes. *Seis. Res. Lett.* **63**, 169–80.
- Oliver, J. (1962). A summary of observed seismic surface wave dispersion. *Bull. Seismol. Soc. Am.* **52**, 81–6.
- Orowan, E. (1960). Mechanism of seismic faulting. *Geol. Soc. Am. Mem.* **79**, 323–45.

- Palmer, A. C. and J. R. Rice (1973). The growth of slip surfaces in the progressive failure of overconsolidated clays. *Proc. R. Soc. A* **332**, 527–48.
- Panza, G. F., F. Vaccari, and R. Cazzaro (1999). Deterministic seismic hazard assessment. In F. Wenzel *et al.* (eds.) *Francea Earthquakes: Tectonic, Hazard and Risk Mitigation*. Springer, Berlin. 269–86.
- Papageorgiou, A. S. and K. Aki (1983). A specific barrier model for quantitative description of inhomogeneous faulting and the prediction of strong ground motion. I, Description of the model. *Bull. Seismol. Soc. Am.* **73**, 693–722.
- Poupinet, G., R. Pillet, and A. Souriau (1983). Possible heterogeneity of the Earth's core deduced from PKIKP travel times. *Nature* **305**, 204–6.
- Prawirodirdjo, L. and Y. Bock (2004). Instantaneous global plate motion model from 12 years of continuous GPS observations. *J. Geophys. Res.* **109**, doi: 10.1029/2003JB002944.
- Press, F. (1956). Determination of crustal structure from phase velocity of Rayleigh waves. I Southern California. *Bull. Geol. Soc. Am.* **67**, 1647–58.
- Prodelh, C. and W. Mooney (2012). *History and Results of Controlled-Source Seismology*. The Geological Society of America, Boulder, 707 pp.
- Rawlinson, N. and M. Sambridge (2003). Seismic travel time tomography of the crust and lithosphere. *Advances in Geophys.* **46**, 81–197.
- Reasenber, P. and D. Oppenheimer (1985). FPFIT, FPLOT, and FPPAGE: FORTRAN computer programs for calculating and displaying earthquake fault-plane solutions, *U.S. Geol. Surv. Open-File Rept.* 85–739, 109 pp.
- Reid, H. F. (1911). The elastic rebound theory of earthquakes. *Bull. Dept Geol. Univ. California* **6**, 412–44.
- Richter, C. F. (1935). An instrument earthquake magnitude scale. *Bull. Seismol. Soc. Am.* **25**, 1–32.
- Ringwood, A. E. (1975). *Composition and Petrology of the Earth's Mantle*. McGraw-Hill, New York.
- Romanowicz, B. (2003). Seismic tomography showing fast anomalies associated with subducted slabs, as revealed by P travel time tomography. *Comptes Rendus Geosciences* **335**, 23–35.
- Romanowicz, B. and B. J. Mitchell (2009). Deep earth structure. Q of the earth from crust to core. In B. Romanowicz and A. Dziewonski (eds.) *Seismology and Structure of the Earth*. Elsevier, Amsterdam, 731–74.
- Romanowicz, B. and D. Dziewonski, eds. (2009) *Seismology and Structure of the Earth*. Elsevier, Amsterdam.
- Satô, Y. (1955). Analysis of dispersed surface waves by means of Fourier transform I. *Bull. Earthq. Res. Inst. Tokyo Univ.* **33**, 33–47.
- Savage, J. C. (1966). Radiation from a realistic model of faulting. *Bull. Seismol. Soc. Am.* **56**, 577–92.
- Savage, J. C. (1972). Relation of corner frequency to fault dimensions. *J. Geophys. Res.* **77**, 3788–95.

- Scholte, J. G. J. (1962). The mechanism at the focus of an earthquake. *Bull. Seismol. Soc. Am.* **52**, 711–21.
- Scholz, C. H. (1982). Scaling laws for large earthquakes: Consequences for physical models. *Bull. Seismol. Soc. Am.* **72**, 1–14.
- Scholz, C. H., L. R. Sykes, and Y. P. Aggarwal (1973). Earthquake prediction: A physical basis. *Science* **181**, 801.
- Schwartz, D. and K. Coppersmith (1984). Fault behaviour and characteristic earthquakes: Examples from the Wasacht and San Andreas fault zones. *J. Geophys. Res.* **89**, 5681–98.
- Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller (2005). High-resolution surface wave tomography from seismic ambient noise. *Science* **307**, 1615–18.
- Shimazaki, K. and T. Nakata (1980). Time predictable recurrence model for large earthquakes. *Geophys. Res. Lett.* **7**, 279–82.
- Sipkin, S. A. (1982). Estimation of earthquake source parameters by the inversion of wave form data. Synthetic wave forms. *Phys. Earth Planet. Inter.* **30**, 242–55.
- Song, X. and D. V. Helmberger (1992). Velocity structure near the inner core boundary from waveform modeling. *J. Geophys. Res.* **97**, 6573–86.
- Sornette, A. and D. Sornette (1989). Self-organized criticality and earthquakes. *Europhys. Lett.* **9**, 197–202.
- Souriau, A. (2009). Deep earth structure: The Earth's cores. In B. Romanowicz and D. Zewoniski (eds.) *Seismology and Structure of the Earth*. Elsevier, Amsterdam, 655–93.
- Sponheuer, W. (1960). *Methoden zur Herdtiefenbestimmung in der Makroseismik*. Akademie Verlag, East Berlin.
- Steketee, J. A. (1958). Some geophysical applications of the theory of dislocations. *Can. J. Phys.* **36**, 1168–98.
- Stoneley, R. (1924). Elastic waves at the surface of separation of two solids. *Proc. R. Soc. A* **106**, 416–28.
- Strelitz, R. A. (1978). Moment tensor inversions and source models. *Geophys. J. R. Astr. Soc.* **52**, 359–64.
- Strelitz, R. A. (1989). Choosing the best double couple from a moment tensor inversion. *Geophys. J. Int.* **99**, 811–15.
- Stump, B. W. and L. R. Johnson (1977). The determination of source properties by the linear inversion of seismograms. *Bull. Seismol. Soc. Am.* **67**, 1489–502.
- Takeuchi, N., R. Geller, and P. Cummins (2000). Complete synthetic seismograms for 3-D heterogeneous Earth models computed using modified DSM operators and their applicability to inversion for Earth structure. *Phys. Earth Plan. Int.* **119**, 25–36.
- Tapley, W. C. and J. E. Tull (1992). *Guide to the UNIX version of SAC*. Livermore (Calif): Lawrence Livermore National Laboratory, University of California.
- Taylor, S. R. and S. M. McLennan (1985). *The Continental Crust: Its Composition and Evolution*. Blackwell, Boston.
- Thomson, W. M. (1950). Transmission of elastic waves through a stratified solid medium. *J. Appl. Phys.* **21**, 89–93.
- Tsuboi, C. (1956). Earthquake energy, earthquake volume, aftershock area and strength of the earth's crust. *J. Phys. Earth* **4**, 63–6.

- Udías, A. (1991). Source mechanism of earthquakes. *Adv. Geophys.* **33**, 81–140.
- Udías, A. and D. Muñoz (1979). The Andalusian earthquake of 25 December 1884. *Tectonophysics* **53**, 291–9.
- Udías, A. and E. Buforn (1988). Single and joint fault-plane solutions from first motion data. In D. Doornbos (ed.). *Seismological Algorithms*. Academic Press, London, pp. 443–53.
- Udías, A. and E. Buforn (1994). Seismotectonics of the Mediterranean region. *Adv. Geophys.* **36**, 121–209.
- Utsu, T. (1961). Statistical study on the occurrence of aftershocks. *Geophys. Mag.* **30**, 521–605.
- Uyeda, S., T. Nagao, and M. Kamogawa (2009). Short-term earthquake prediction: Current status of seismo-electromagnetics. *Tectonophysics* **470**, 205–13.
- Vaněk, J., A. Zápotek, V. Karník, *et al.* (1962). Standardization of magnitude scales. *Izv. Akad. Nauka SSSR, Geofiz.* **2**, 153–8.
- Varotsos, P. and M. Lazaridou (1991). Latest aspects of earthquake prediction in Greece based on seismic electric signals. *Tectonophysics* **188**, 321–47.
- Von Hoff, K. A. (1840). *Chronik der Erdbeben und Vulkan-Ausbrüche*, Justus Perthes, Gotha, Germany.
- Vvedenskaya, A. V. (1956). Determination of displacement fields for earthquakes by means of the dislocation theory. *Izv. Akad. Nauka SSSR, Geofiz.* **3**, 277–84 (in Russian).
- Waldhauser, F. and W. L. Ellsworth. (2000). A double-difference earthquake location algorithm: method and application to the Northern Hayward Fault. *Bull. Seismol. Soc. Am.* **90**, 1353–68.
- Wickens, A. J. and J. H. Hodgson (1967). Computer re-evaluation of earthquake mechanism solutions (1922–1962). *Publ. Dom. Obs. (Ottawa)* **33**, 1–560.
- Wieland, E. and G. Streckeisen (1982). The Leaf-Spring Seismometer: Design and Performance. *Bull. Seismol. Soc. Am.*, **72**, 2349–67.
- Wilson, J. T. (1965). A new class of faults and their bearing on continental drift. *Nature* **207**, 343.
- Woodhouse, J., D. Giardini, and X. Li (1986). Evidence for inner core anisotropy from free oscillations. *Geophys. Res. Lett.* **13**, 1549–52.
- Woodward, R. L. and G. Master (1991). Global upper mantle structure from long period differential travel times. *J. Geophys. Res.* **96**, 6351–77.
- Worthington, L., K. C. Miller, E. A. Erslev, *et al.* (2016). Crustal structure of the Bighorn Mountains region: Precambrian influence on Laramide shortening and uplift in north-central Wyoming. *Tectonics* **35**, 208–36.
- Wu, Y.-M. and H. Kanamori (2005). Experiment on an onsite early warning method for the Taiwan early warning system, *Bull. Seismol. Soc. Am.* **95**, 347–53.
- Wyss, M. and D. Booth (1997). The IASPEI procedure for evaluation of earthquakes precursors. *Geophys. J. Int.* **131**, 423–4.
- Young, C. J. and T. Lay (1987). The core–mantle boundary. *Ann. Rev. Earth Planet. Sci.* **15**, 25–46.
- Zahn, J. (1696). *Specula physico mathematica (Scrutinium IV Geocosmicum. Cap. XIII, Chronicon)*, Johannis Christopher Lochner, Nuremberg, Germany.

- Zeilinga de Boer, J. and D. T. Sanders (2005). *Earthquakes in Human History*. Princeton University Press, Princeton. 278 pp.
- Zhou, H.W. (2014). *Practical Seismic Data Analysis*. Cambridge University Press, Cambridge.
- Zuccolo, E., F. Vaccari, A. Peresan, and G.F. Panza (2011). Neo-deterministic and probabilistic seismic hazard assessments. A comparison over the Italian territory. *Pure Appl. Geophys.* **168**, 69–83.
- Zürn, W., A. M. G. Ferreira, R. Widmer-Schmidrig, *et al.* (2015). High-quality lowest-frequency normal mode strain observations at the Black Forest Observatory (SW-Germany) and comparison with horizontal broad-band seismometer data and synthetics. *Geophys. J. Int.* **203**, 1786–1803.

Index

Italic entries are key terms

- acceleration, 37, 54, 94, 453
- accelerograms, 54
- accelerographs, 54
- Adams-Williamson, 237
- adiabatic conditions, 71
- aftershocks, 22, 452
- Airy's phase*, 291
- ak135 model, 229
- aliasing*, 515
- analogical records*, 45
- anelasticity, 347
- angle of incidence*, 101
- angle of polarization*, 102
- anisotropic media, 351
- anisotropy, 335
 - in the Earth, 335
- anthropogenic, 477
- apparent angle
 - incidence, 141
 - polarization, 141
- apparent velocity, 147, 168
- arrays, 55
- aseismic slip, 482
- asperities*, 451, 452
- asthenosphere*, 26, 304, 350
- attenuation*
 - body waves, 343
 - coda waves, 348
 - free oscillations, 346
 - harmonic damped, 338
 - surface waves, 345
- average stress*, 374
- azimuth*, 101
- azimuthal anisotropy, 355
- b parameter, 484
- barriers*, 451, 452
- Benioff seismograph, 34
- Benndorf relation, 168
- Bessel equations*, 107, 506
- Bessel functions*, 107, 319, 506
 - spherical, 107
- Betti's reciprocity theorem*, 82
- bilateral fractures*, 434
- body forces*, 74, 88, 382, 415
- body waves, 24, 88
- Brune's model*, 438
- bulk modulus*, 66
 - adiabatic, 69
 - isothermal, 69
- Byerly's method*, 458
- Cauchy tensor*, 61
- cellular automata, 487
- centroid*, 360
- centroid moment tensor, 427
- characteristic earthquake, 496
- circular fault, 375
 - Brune's model, 438
 - dynamic's model, 448
 - Savage's model, 440
- coda waves, 348
 - attenuation, 357
- cohesive zone*, 449
- Conrad discontinuity, 25
- conservation of energy*, 73
- conservation of mass*, 72
- constructive interference, 259, 266, 278, 282
- continuity equation, 96
- convolution, 35, 82, 385
- core, 25
 - composition, 256
 - density, 256
 - inner, 251
 - outer, 251
 - phases, 251
 - transition zones, 256
 - velocity models, 229, 251
- core-mantle boundary, CMB, 25
- corner frequency*, 436, 440, 454, 470
- Coulomb's stress transfer, 481
- critical incidence, 120
- crust*, 24
 - continental, 25, 27, 239
 - mineral composition, 244
 - oceanic, 27, 241, 244
 - orogenic, 242
 - structure, 242
 - wave phases, 238

- cut-off frequency, 268
- cylindrical anisotropy, 356
- cylindrical coordinates, 110, 280, 502
- cylindrical waves*, 109
- damping, 33, 35, 335
 - critical, 339
 - factor, 36
- deep earthquakes, 20, 246, 460, 477, 488
- deep foci, 200
- delay time, 172, 179
- density distribution, 237
- deviatoric stress*, 65
- diffracted waves, 253
- digital seismographs, 45
- dilatancy, 495
- dipping layer, 176, 231
- directivity function*, 437
- dislocation, 388, 418, 456
 - shear, 400, 408, 432
- dispersed waves, 269, 289, 292, 345
- dispersion, 96, 271
 - curves, 268, 272
 - Earth structure, 300
 - equation, 266
 - Love, 279
 - Rayleigh, 277
- displacement-stress, 155
- double couple, 386, 422, 459
- dynamic models*, 382, 444
- dynamic range, 39
- Earth's structure, 14, 228
- earthquake early warning systems, 5, 30, 493
- earthquake engineering, 6, 54
- earthquakes
 - catalogue, 2, 359
 - doublets, 481
 - historical, 5, 360, 483
 - interplate, 376
 - intraplate, 376
 - precursor, 5, 494
 - spatial distribution, 20, 477
 - temporal distribution, 22, 480
- eikonal equation, 161
- elastic coefficients*, 65
- elastic energy, 70, 341, 389, 484
 - flux, 446
 - intensity, 123, 205, 224
 - partition, 123, 136
 - potential, 70, 71
- elastic rebound*, 380
- elastic rod, 313
 - longitudinal vibrations, 313
 - torsional vibrations, 315
- elastic string, 308, 310
- elastic waves, 23, 87, 115
- elasticity, 60
- empirical Green's functions, 464
- energy
 - seismic, 373, 374, 444
- energy magnitude relations, 373
- energy propagation, 96, 97, 224
- epicenter, 167, 358
 - instrumental determination, 361
 - macroseismic determination, 359
- epicentral distance, 35
- equation of motion, 73
 - seismograph, 36
- equivalent forces*, 382
- Euler's equation*, 77, 95
- excitation functions*, 389, 468
- explosive source, 418
- far field, 396, 424, 432
- fault, 6, 18
 - area, 360
 - normal, 381
 - orientation, 381
 - plane, 374
 - reverse, 381
 - strike. *See azimuth*
 - strike-slip, 381
 - types, 381
 - vertical, 381
- fault creep*, 20
- fault-plane solutions*, 458
 - graphical methods, 460
 - numerical methods, 461
- focal depth, 249, 358, 466
 - graphical methods, 362
 - intensity, 367
- focal sphere, 458
- foreshock*, 23, 452
- Fourier analysis of seismograms*, 297
- Fourier transform*, 511
- fractal, 486
- fracture, 2, 18, 380, 388
 - complexity, 449
 - dynamic model, 431
 - geometry, 381
 - kinematic model, 431
 - shear, 419
 - shear point source, 400
- free oscillations, 308
 - elastic sphere, 322
 - liquid sphere, 318
 - observations, 331
 - spheroidal modes, 323
 - toroidal modes, 323
- free surface motion
 - at the liquid, 126
 - at the solid, 139
- friction models, 448
- fundamental mode
 - free oscillations, 309

- fundamental mode (cont.)
 - guided waves, 266
 - Love waves, 274
 - Rayleigh waves, 278
- Galitzin seismograph, 33
- generalized inverse matrix, 153
- geometric spreading, 107
- GPS, 5, 56, 483
- Green's function, 83, 384
 - far field, 397
 - infinite medium, 390
 - near field, 397
- Green–Volterra formula, 82
- Griffith's energy, 446
- ground motion, 29, 35
- group velocity, 96
 - determination, 294
- Gutenberg and Richter
 - law, 484
 - magnitude, 369
- harmonic motion*
 - damped, 37
- harmonics, 325
- Haskell's model, 434
- head wave, 121, 171, 177
- healing front, 443
- Helmholtz
 - equation, 89
 - function, 71
 - theorem, 78
- Herglotz–Wiechert
 - plane medium, 203
 - spherical medium, 223
- Hooke's law, 65
- Huygens' principle, 127
- hypocenter, 12, 167, 358
 - determination, 360
- iasp91 model, 229
- IASPEI, 8
- impulsive
 - acceleration, 38
 - force, 79, 390
 - time function, 285
- incidence. See angle of incidence*
 - normal, 120
- inhomogeneous waves, 121
- inner core, 25, 251
- instrumental response, 35, 49
- intensity, 17
 - attenuation, 367
 - instrumental, 367
 - maps, 359, 367
 - scales, 365
- inverse problem, 7, 202, 380
- isoseismal maps. See intensity maps*
- isothermal bulk modulus, 69
- isotropic, 6, 66, 335
- isotropy transverse, 352
- Jeffreys–Bullen tables, 228
- kinematic models*, 382
- Lagrangian formulation, 77, 187
- Lamé's coefficients, 66
- Laplace transform, 511
- Legendre functions, 508
- lithosphere, 26, 245
- Love waves, 269, 271
 - dispersion equation, 272
 - displacements, 274
 - modes, 274
- low-velocity, 27, 179, 199, 221, 239, 246
- magnification, 33, 45
 - curve, 39
 - dynamic, 39
- magnitude*, 17
 - distribution, 483
 - saturation, 372
 - scales, 368
- main shock*, 22
- mantle*, 24, 237
 - composition, 251
 - lower, 247
 - upper, 244
- Maxwell body*, 336
- microseisms*, 51
- migration
 - earthquakes, 23
 - technique, 233
- mode, 266
 - longitudinal, 315
 - spheroidal, 327
 - toroidal, 327
 - torsional, 316
- Moho
 - discontinuity, 24
 - phases, 238
- Mohorovičić's law, 217
- moment tensor*
 - inversion, 427
 - separation, 422
 - synchronous, 424
- multiplets, 331
- Navier's equation*, 76
- near field*, 367, 397, 453
- network, seismologic, 34, 305

- Newton's second law, 36, 73
 noise, 51
 Nyquist frequency, 515

Ocean Bottom Seismographs (OBS), 56
 oceanic ridges, 21, 477
 Omori's law, 480
 origin time, 12, 358
 Orowan's condition, 378
 outer core, 25
 overtones, 309

 P wave, 24, 88
 displacements, 99, 101
 paleoseismicity, 5, 483
 particle motion, 143, 278
 phase velocity, 91, 96, 285
 apparent, 329
 determination, 296
 plane waves, 98, 266
 plate boundaries, 27, 477
 plate tectonics, 19, 26, 245
 point source, 106, 204, 382, 414
 dislocation, 400
 displacements, 424
 isotropic, 358
 Poisson
 distribution, 22, 480
 equation, 97
 ratio, 67
 solid, 68
 polarization angle, 102
 poles and zeros, 46
 potential. *See elastic energy*
 potential function, 71
 of displacements, 78
 of forces, 79
 prediction, 5, 30, 481
 Preliminary Earth Model (PREM), 229
 profile
 dipping layer, 178
 reflection, 230
 refraction, 174, 230
 propagator matrix, 150
 general problem, 157
 P-SV motion, 158
 SH motion, 154

 quality factor (Q), 341
 observations, 356
 P and S waves, 343

 radiation pattern, 396
 Green functions, 398
 shear fracture, 401
 radius of curvature, 184, 191, 213, 217
 radon gas, 495

ramp function, 397
 ray beams, 204, 224
 ray parameter
 plane, 116
 plane medium, 184
 spherical medium, 209
 ray path (trajectories), 98, 165
 variable velocity, 184
 ray theory, 161
 condition of validity, 164
 variable velocity, 183
 ray tracing, 180
 Rayleigh waves
 dispersion, 277
 displacements, 263
 half-space, 259
 layer over half space, 276
 Rayleigh's equation, 262
 receiver functions, 235
 reduced times. *See tau functions*
 reflection
 coefficients, 118
 elastic media, 127
 free surface, 132
 two liquid media, 116
 refraction
 coefficients, 118
 elastic media, 127
 two liquid media, 116
 representation theorem, 83
 resilience, 16
 rift zones, 27, 487
 rigidity modulus. *See shear modulus*
 rise time, 406
 rupture, 358, 427
 arrest, 440
 nucleation, 440
 propagation, 440
 velocity, 440

 S wave, 24, 88
 displacements, 99, 101
 Savage's model, 440
 SEED format, 47
 seismic cycle, 481
 seismic gap, 494
 seismic hazard, 29, 490
 deterministic assessment, 490
 probabilistic assessment, 491
 seismic moment, 18, 374, 410
 determination, 470
 scalar, 372
 tensor, 414
 seismic risk, 29, 490
 seismic tomography, 233
 seismic zonation, 491
 seismicity, 4, 20, 477

- seismogenic layer*, 18, 381, 486
- seismogenic zone, 491
- seismograms, 4, 32, 35
- seismographs, 32
 - broad-band, 32, 35
 - digital. *See digital seismographs*
 - electromagnetic, 40
 - mechanical, 4
- seismology, 1
 - division, 5
 - multidisciplinary science, 4
- seismometer
 - theory, 35
- seismometer galvanometer, 33
- seismometry, 6
- seismoscope, 33
- seismotectonics, 4, 26, 487
- self-organized criticality, 495
- self-similarity, 486
- shadow zone*, 199, 221, 244, 248
- shear modulus*, 66
- single couple, 385
- site effects, 54, 55, 359
- slip, 374, 388, 418
 - angle, 381
 - distribution, 471
 - rate, 401
 - velocity, 389
 - weakening, 450
- slowness*, 147, 186
- Snell's law, 115, 166, 209
- Somigliana's tensor*, 81
- source dimensions, 432
- source mechanism, 380
- source time function*, 406
- spectra*, 50
 - amplitude, 433
 - free oscillations, 331
 - scalar seismic moment, 372
- spherical waves*, 105
- splitting, 331
- stack technique, 232, 331
- standing waves, 309
- static problem, 444
- stationary phase*, 289
- stick-slip, 485
- Stoneley waves, 281
- strain tensor, 62
- stress drop*, 374, 444
 - dynamic, 448
 - static, 448
- stress tensor, 65
- strong motion*, 47
 - instruments, 54
- Sturm–Liouville equation, 317
- subduction zones*, 27
- surface waves, 24, 259
- swarm of earthquakes*, 23
- take-off angle, 168, 224, 405
- tau function, 169
- tectonosphere, 246
- time residuals
 - hypocenter, 363
 - tomography, 233
- time-predictable model, 481
- total reflection, 121
- transcurrent boundary, 27
- transfer function*, 39
- transform faults*, 27
- travel time, 167
 - deep focus, 200
 - dipping layer, 176
 - Earth structure, 228
 - plane medium constant velocity, 167
 - plane medium variable velocity, 185
 - sphere constant velocity, 212
 - sphere variable velocity, 214
- tsunami*, 14
- unilateral fractures*, 434
- upper mantle, 25
- velocity
 - apparent, 116
 - energy transport, 97
 - rupture, 431
- viscous body*, 336
- Wadati–Benioff zone*, 28
- wave equation*, 87
 - solutions, 89
- wave propagation in anisotropic media, 352
- wave-form modeling, 463
- websites, 9
- work elastic, 70
- Young's modulus*, 66
- Zöppritz's equations, 131